



OPEN An integrated approach of support vector machine (SVM) and weight of evidence (WOE) techniques to map groundwater potential and assess water quality

Malik Talha Riaz^{1,2}, Muhammad Tayyib Riaz², Adnanul Rehman³, Ahmed Ali Bindajam⁴, Javed Mallick⁵ & Hazem Ghassan Abdo⁶✉

This study addresses the critical need for effective groundwater (GW) management in Muzaffarabad, Pakistan, amidst challenges posed by rapid urbanization and population growth. By integrating Support Vector Machine (SVM) and Weight of Evidence (WOE) techniques, this study aimed to delineate GW potential zones and assess water quality. This study fills the gap in applying advanced machine learning and geostatistical methods for accurate GW potential mapping. Eight thematic layers based on topography, hydrology, geology, and ecology were utilized to compute the GW potential model. Additionally, water quality analysis was performed on collected samples. The findings indicate that flat and gently sloping terrains, areas with an elevation range of 611–687 m, and concave slope geometries are associated with higher GW potential. Additionally, proximity to drainage and high-density lineament zones contribute to increased GW potential. The results showed that 31.1% of the area had excellent GW potential according to the WOE model, whereas the SVM model indicated that only 20.3% fell in the excellent potential zone. Results showed that both models performed well in the delineating GW potential zones. Nevertheless, the application of the SVM method is highly recommended which will be benefited in GW resources management related to urban planning. The study also evaluates the spatial distribution of GW quality, with a focus on physical and chemical parameters, including electrical conductivity, pH, turbidity, total dissolved solids, calcium, magnesium, chloride, nitrate, and sulphate. Bacterial contamination assessment reveals that 76% of spring water samples (30 out of 39 samples) are contaminated with *E.coli*, raising public health concerns. Based on the chemical analysis of GW samples the study identified exceedances of WHO guidelines for calcium in two samples, magnesium in seven samples, sulphate in ten samples, and nitrate levels were below the WHO guideline across all samples. These results highlight localized chemical contamination issues that require targeted remediation efforts to safeguard water quality for public health.

Keywords GW potential, GW quality, Muzaffarabad, WOE, SVM, Sustainable GW management

Water, the lifeblood of our planet, has been essential for sustaining all forms of life since the beginning of time. It is a vital resource that nurtures ecosystems and supports human existence. However, the quality and availability of water have become significant global concerns in recent years, primarily because of climate change and population growth¹. Access to clean water is a pressing issue worldwide, particularly in underdeveloped countries. Recognizing this challenge, the Millennium Development Goals were established to improve access

¹Institute of Earth Science, University of Silesia, International Environmental Doctoral School, ul. Będzińska 60, Sosnowiec 41-200, Poland. ²Institute of Geology, King Abdullah Campus, University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan. ³College of Geography and Environmental Science/Key Research Institute of Yellow River Civilization and Sustainable Development and Collaborative Innovation Center on Yellow River Civilization of Henan Province, Henan University, Kaifeng 475004, PR China. ⁴Department of Architecture, College of Architecture and Planning, King Khalid University, Abha 61411, Saudi Arabia. ⁵Department of Civil Engineering, College of Engineering, King Khalid University, Abha, Kingdom of Saudi Arabia. ⁶Geography Department, Faculty of Arts and Humanities, Tartous University, Tartous, Syria. ✉email: Hazemabdo@tartous-univ.edu.sy

to safe drinking water sources for communities around the globe². By focusing on this fundamental aspect of human well-being, we can address one of the most basic yet essential needs of individuals and empower communities to thrive. However, managing water resources effectively has become increasingly challenging in the face of escalating human activities. Activities such as the discharge of wastewater and sewage into surface and groundwater (GW) resources pose significant threats to water quality and overall ecosystem health³. The consequences of these actions extend far beyond human well-being and impact the delicate balance of our planet's biodiversity.

Furthermore, rapid urbanization and population growth have placed immense pressure on GW resources, leading to their overuse and subsequent depletion⁴. Aquifers, the underground reservoirs that provide a significant portion of our freshwater supply, are at risk of irreversible decline. This highlights the urgent need for sustainable water management practices and the development of innovative solutions to ensure the long-term viability of this finite resource⁵. In recent decades, the combination of climate variability, rapid population growth, and overexploitation has detrimental effects on valuable GW resources worldwide^{6–9}. Climate variability, including changes in rainfall patterns, surface runoff, evaporation, and rising temperatures, has further exacerbated the issue¹⁰. Despite these challenges, GW remains a crucial resource, with subsurface water reserves exceeding those found in streams and freshwater lakes by over 100 times.

Countries facing water scarcity and impacted by climate change-related changes in surface and GW quantity and quality encounter difficulties in understanding the dynamics of their GW resources¹¹. This is particularly true for countries like Pakistan, where rapid GW use, especially in agriculture-dependent regions, has become a pressing concern for decision-makers responsible for managing sustainable GW resources. As a result, comprehending the GW resource regime has become a critical task for countries grappling with water shortages and climate change impacts¹¹. Pakistan is among the most water-stressed countries in the world, and the growing GW quality problems are of grave concern for the increasing population^{12,13}. Water pollution due to industrial, agricultural, and domestic sources has led to widespread contamination, affecting approximately 80% of the population with access to unsafe drinking water¹⁴. Urgent measures are needed to address these issues to ensure sustainable water management and achieve water-related UNSDGs by 2030¹⁵. In recent years, the use of remote sensing (RS) and Geographic Information Systems (GIS) has proven effective in identifying potential GW zones^{16–18}. RS provides valuable information and a comprehensive view of spatial and temporal distribution, enabling a more efficient assessment of regional GW flow dynamics¹⁹. The integration of RS science, GIS, and ground-based field research is crucial for predicting GW dynamics. Various factors such as geology, drainage patterns, slope, elevation, land use/land cover, lineament density, soil type, rainfall, topography, lithology, porosity, and climatic conditions play a significant role in understanding the occurrence, movement, and delineation of potential GW zones²⁰. Various statistical and machine learning (ML) models have been developed and employed for GW prediction. Models such as the Analytical Hierarchy Process (AHP)^{21,22}, Weight of Evidence-WOE²³, Artificial Neural Networks^{24,25}, Support Vector Machine-SVM^{26,27}, Adaptive Network-based Fuzzy Inference System-ANFIS²⁸, Extreme Gradient Boosting-XGB²⁹, Fuzzy Logic³⁰, and the M5 model have been widely adopted for GW prediction. Choubin and Rahmati³¹ adopted Random Forest (RF) method for GW potential mapping in fractured bedrock aquifer in Iran. Choubin, et al.³² successfully generated a GW potential map for the Firoozeh watershed in Iran using the Classification and Regression Trees (CARTs) algorithm. Their study incorporated eleven conditioning factors, achieving a notable accuracy level of 88%. Mosavi, et al.³³ conducted a study to evaluate four ensemble models (GamBoost, AdaBoost, Bagged CART, and RF) for GW mapping using 339 GW resource locations and spatial conditioning factors. Their findings highlight that Bagging models, particularly RF, demonstrated superior performance (accuracy=0.86), with topographic and hydrological variables playing crucial roles in the modeling process. Recently Saha, et al.³⁴ employed a ML and geospatial data integration approach to manage urban and peri-urban aquifer sustainability in Vizianagaram, Southern India. Using RF and SVM models with hydrogeological and geo-environmental data, they categorized GW potential with prediction accuracies exceeding 80%, notably 88.40% for RF.

In our study, we chose to employ SVM and WOE techniques for several compelling reasons that align with the specific objectives and characteristics of our research area. SVM was preferred due to its ability to handle complex, nonlinear relationships inherent in GW potential mapping, offering high accuracy and robustness in classification tasks³⁵. This method excels in effectively utilizing high-dimensional data and generalizes well to predict GW potential in unobserved areas without significant risk of overfitting, which is crucial for spatial predictions. SVM has gained popularity in hydrology, as evidenced by their increasing adoption^{26,29,36}. Additionally, WOE was selected for its simplicity and interpretability, allowing for the transparent integration of diverse thematic layers based on their statistical relevance to GW occurrence^{37,38}. This approach facilitates a clear understanding and communication of results to stakeholders, essential for informed decision-making in GW management.

Water-shortage countries, particularly those affected by climate change impacts on surface and GW quantity and quality, face challenges in comprehending the dynamics of their GW resources. The rapid expansion of GW use, especially in agrarian-based developing nations like Pakistan, raises concerns for decision-makers responsible for managing sustainable GW resources. Monitoring GW observations from natural springs and wells is essential to identify stress zones losing GW storage and assess GW quality in the area. A comprehensive study of GW storage is vital for understanding the hydrological cycle and its relationship with climate change¹¹. In the Northwestern Himalayas, where cities are located along river banks, water quality is adversely affected by rapid population growth³⁹. Pakistan has a significant portion of the population lacking access to safe drinking water, below the standards set by the World Health Organization (WHO). In northern Pakistan, the population relies on natural springs and wells for drinking water, which also serve as irrigation sources for nearby agricultural areas. In Pakistan Administered Kashmir (PAK), over 80% of illnesses have been attributed to the consumption of poor-quality water from both surface and GW sources⁴⁰. Proper management of GW quality is crucial for

protecting these resources from contamination¹¹. The Muzaffarabad municipality is facing challenges related to GW management and water quality⁴¹. The rapid population growth, coupled with limited access to safe drinking water, highlights the need for effective GW potential mapping and water quality assessment. Therefore, there is a pressing need to address these issues and develop sustainable strategies for GW resource management in the Muzaffarabad municipality. There is a significant research gap in the study area, regarding baseline studies for GW potential mapping and quality assessment. Despite the critical importance of sustainable water management in the region, there is a lack of comprehensive studies addressing these aspects. This study aims to fill this gap by systematically mapping GW potential and assessing water quality in Muzaffarabad, providing essential data for informed decision-making and sustainable water resource management. The research objectives encompass mapping GW potential using SVM and WOE, assessing water quality through geochemical tests, evaluating the effectiveness of the mapping techniques, identifying factors influencing GW potential and water quality, and providing recommendations for sustainable GW resource management. The novelty of this research lies in the integration of SVM and WOE techniques for GW potential mapping and the comprehensive assessment of water quality in the specific context of the Muzaffarabad municipality. We integrated the two methods in a complementary way, not in a real integration. By complementary, we mean that we used the strengths of each method to compensate for the weaknesses of the other. The findings of this study will contribute valuable insights into the field of GW resource management and serve as a basis for sustainable strategies in the municipality.

Study area characteristics

The study area is geographically situated in the northeastern Himalayas of Pakistan and is located 34°20'00" to 34°25'00" N latitude and 73°26'00" to 73°30'45" E longitude covering an area of 52.1 km² (Fig. 1). Topographically, the study area is situated on the conflux of the Neelum and Jhelum rivers having rugged topography dissected by several other minor and major streams. It is a hilly mountainous area with its elevation varying from 611 m to 1,630 m above mean sea level (MSL). Muzaffarabad, with a population of 149,005, faces challenges in GW availability due to its rugged terrain, particularly away from streams. The region's aquifers, predominantly unconfined and situated near the Neelum River, rely on recharge from seasonal rainfall and the river itself⁴². The area experiences an average annual rainfall of 1400 mm, with temperatures ranging from 18 °C to 39 °C in winter and summer respectively. The local ecosystem is under significant pressure due to high population density, highlighting the critical link between GW potential and human sustainability⁴³. Lithologically study area consists of different rock units of sedimentary rocks. These include the Pre-Cambrian Hazara Formation the Cambrian Muzaffarabad Formation, the Paleocene-Eocene sequence, Miocene Murree Formation, and the Quaternary Alluvium deposits (Fig. 2). The study area has severe challenges regarding the GW potential and quality. The rugged topography, combined with the presence of numerous minor and major streams, obstruct the availability and accessibility of GW resources. This challenging terrain and complex hydrogeological conditions make it imperative to establish a dependable method for identifying viable GW sources in the research area. Addressing these issues is of utmost importance, as the scarcity of reliable GW sources adversely impacts the well-being of local inhabitants and hinders sustainable development in the region. Therefore, conducting a comprehensive study to determine the GW potential and availability is crucial to developing effective strategies for sustainable water management and ensuring the prosperity of the local community in the study area.

Methodology

The GW potential mapping and water quality assessment methodology for the Muzaffarabad municipality consists of several key steps (Fig. 3). It begins with the collection of relevant data, including geological information, drainage patterns, slope, elevation, land use/land cover, and lineament density. Additionally, water quality data from spring waters and well bores is obtained from Environmental Protection Agency (EPA), AJK, Pakistan. To ensure accuracy, the collected data undergoes a thorough preprocessing stage, which involves cleaning the data and applying normalization or standardization techniques for uniformity and comparability. The cleaning process primarily focused on removing outliers and correcting any errors, particularly those originating from areas outside the study area. For normalization, we applied the min-max approach to ensure uniformity and comparability in the dataset. This technique scales the data to a specific range (typically 0 to 1), providing a standardized format for the variables. For GW potential mapping using SVM, the dataset is divided into training and testing sets. To address potential biases in the model, we implemented a strategy of randomization of data. The dataset was randomly shuffled before splitting into 70% training and 30% testing sets. This ensures that the distribution of data points is uniform across both sets, minimizing the risk of any inherent biases affecting the model's performance. A total of 39 samples were collected for the study. The data was split in a 70:30 ratio, with 70% of the samples (27 samples) allocated for training purposes, and the remaining 30% (12 samples) reserved for testing the model. The SVM algorithm is then applied to the training dataset to construct a predictive model. During this process, an appropriate kernel function is selected, and hyperparameters are fine-tuned to optimize the model's performance. The accuracy and effectiveness of the SVM model are evaluated and validated using the testing dataset. To evaluate and validate the SVM model, we used some metrics (Accuracy, F1-score and AUC, the area under the ROC curve) that can measure its performance and accuracy. Once validated, the trained SVM model is extended to cover the entire study area, generating a comprehensive GW potential map. Simultaneously, GW potential mapping is also conducted using the WOE technique. WOE values are calculated for each predictor variable based on their relationship with GW potential. These WOE values are then weighted to determine their relative importance in predicting GW potential. The weighted WOE values are combined to calculate an overall WOE score for each location within the study area. As a result, the study area is classified into different GW potential zones, leading to the creation of a GW potential map. Alongside GW potential mapping,

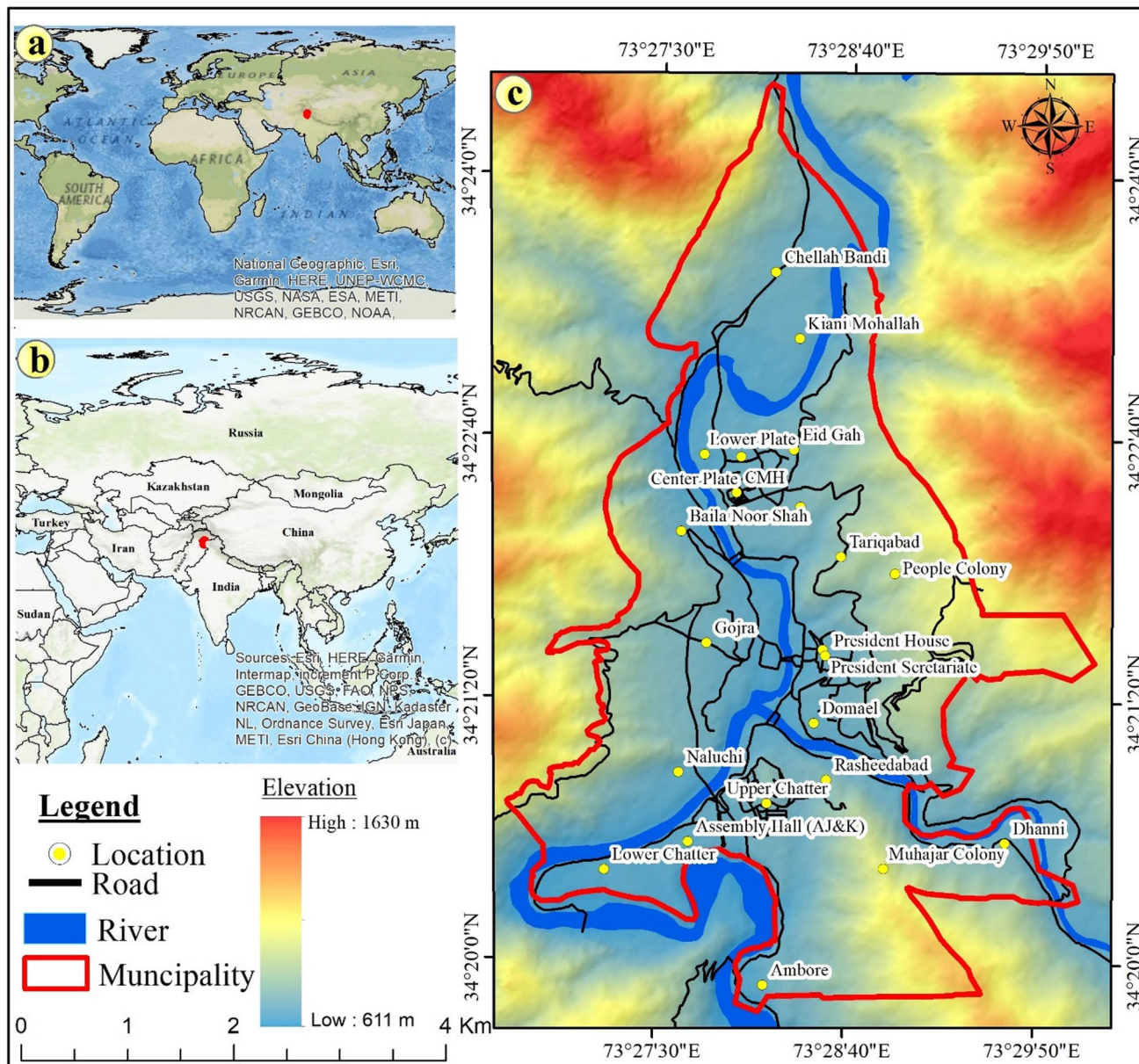


Figure 1. Geographic location of area (a) Geographic location of the area with respect to world map (b) location of the study area in NE part of Pakistan, (c) location map: visualizing the study area's geographical location. (Source: Figures a and b were generated using ESRI online resources).

a water quality index is also developed using the collected water quality data. Various geochemical parameters, including the physical and chemical properties of water, are analyzed following the standard procedures outlined by the WHO. The water quality index is calculated based on these parameters to assess and evaluate the overall water quality at different locations within the municipality. In summary, this methodology efficiently combines the SVM and WOE techniques for GW potential mapping and integrates geochemical tests for water quality assessment. This combined approach enables the creation of GW potential maps and water quality index maps for the Muzaffarabad municipality, ensuring a consistent and comprehensive analysis.

Dataset used

In this study, the GW potential model was computed using eight thematic layers based on topographical, hydrological, geological, and ecological parameters. The topographical factors, including elevation, slope gradient, aspect, and curvature, were derived from the Phased Array type L-band Synthetic Aperture Radar (PALSAR) Digital Elevation Model (DEM) with a 12.5 m resolution obtained from the Advanced Land Observing Satellite (ALOS). The Spatial Analyst tool of ArcMap 10.8 was used for this purpose.

An area's topographic characteristics play a crucial role in determining the water table elevation². Additionally, geographical factors such as surface drainage and stream networks are significantly influenced

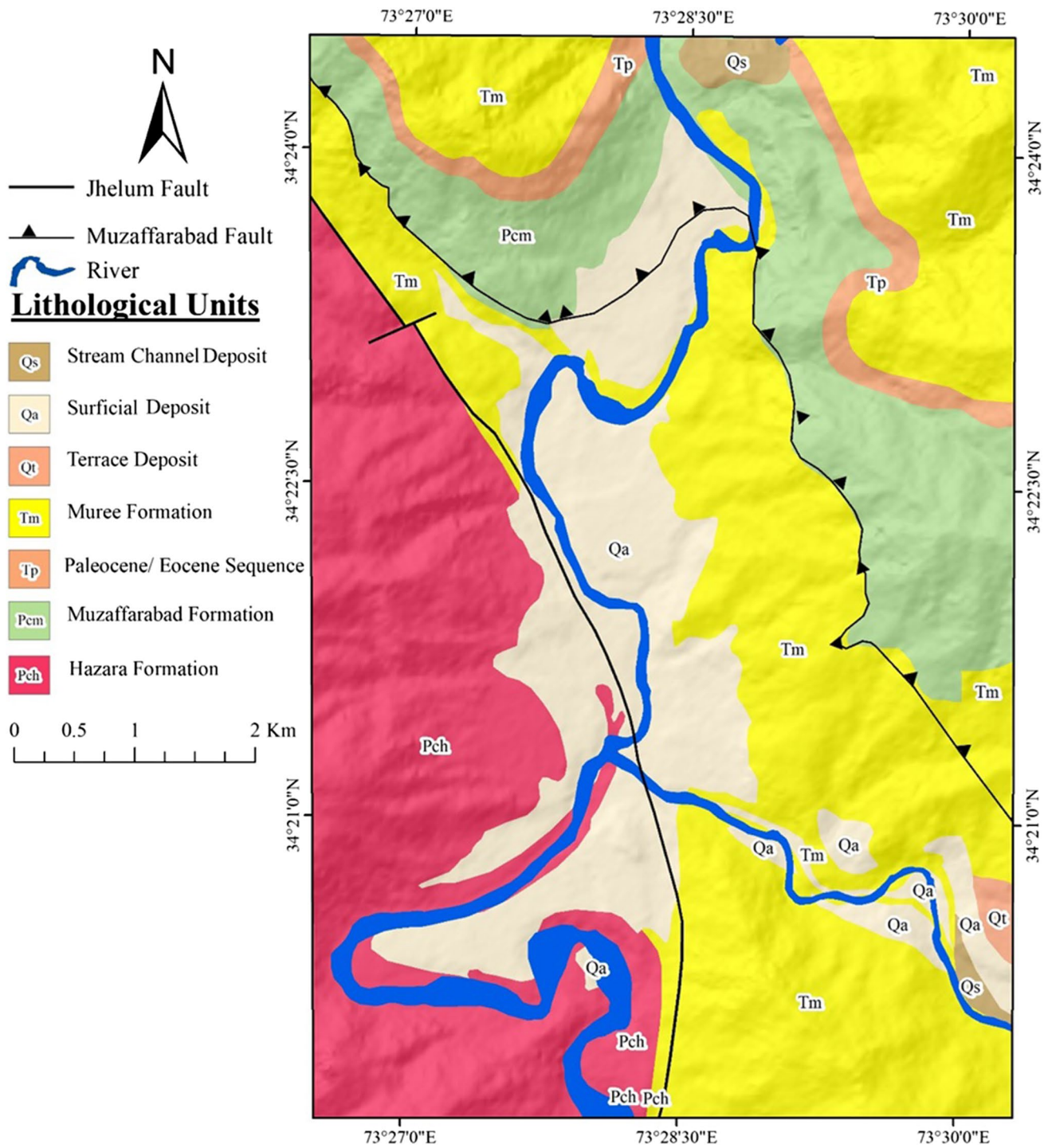


Figure 2. Geological map of the study area (digitized from³⁷).

by elevation. The maximum elevation in the area was 1630 m, which was divided into five classes (Fig. 4a). The decision to categorize elevation into five classes was based on a combination of geological and hydrogeological considerations, as well as practical considerations for the GW potential mapping study. The slope gradient indicates the degree of inclination or steepness of a surface and is relevant for determining runoff. The slope of the research area was classified into six classes: very mild slope (0–10), mild slope (11–19), gentle slope (20–26), moderate slope (27–33), steep slope (34–46), and extremely steep slopes (above 46) (Fig. 4b). The slope aspect, representing the direction a slope faces, is also essential as it affects snow melting and water infiltration. The slope aspect of the study area was classified into eight classes (Fig. 4c). Curvature, indicating the nature of the surface profile as concave or convex, influences water accumulation in the study area. The curvature of the slope was classified into three classes (Fig. 5a). Hydrological parameters, such as drainage patterns, are significant

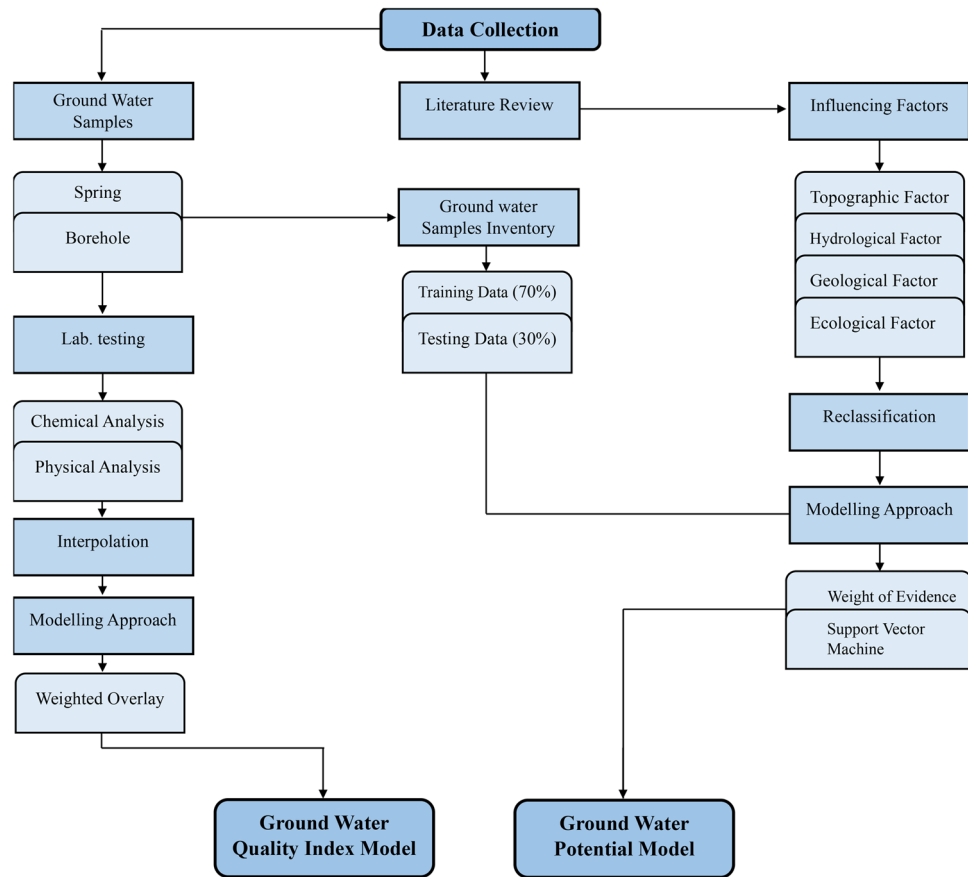


Figure 3. Methodological flow diagram.

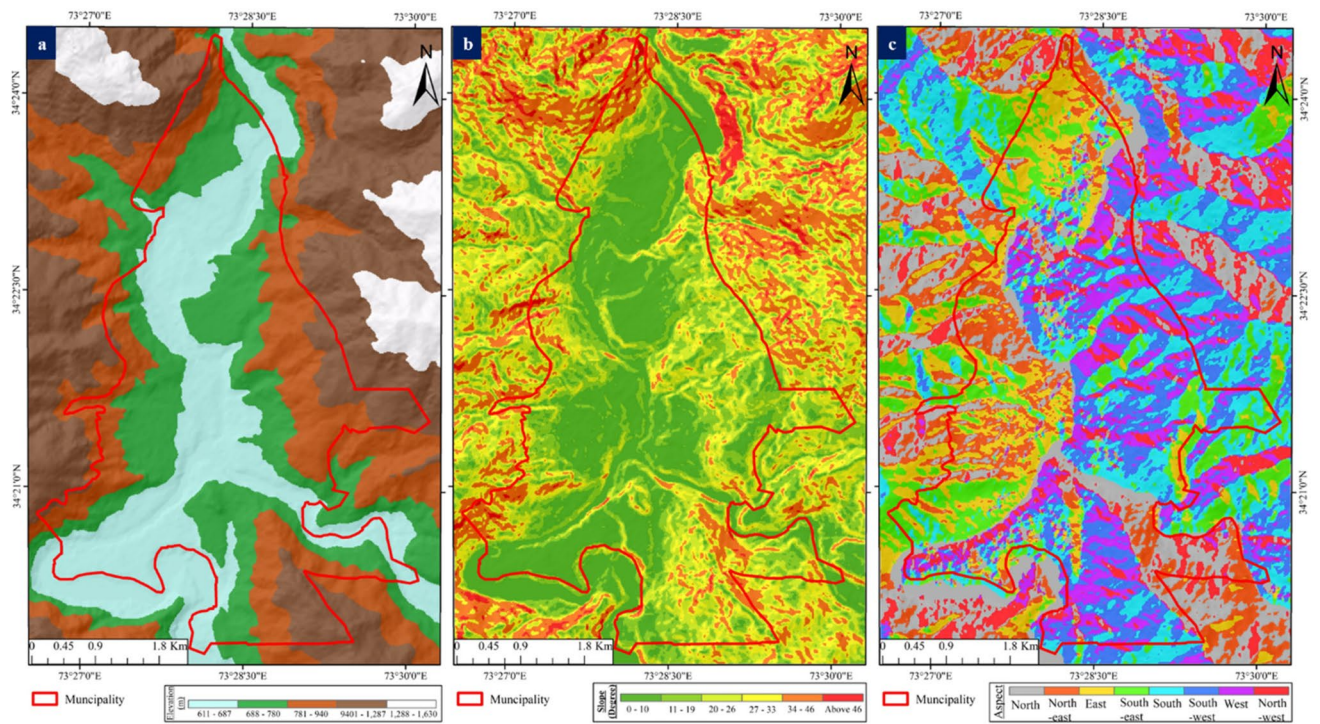


Figure 4. Groundwater potential influencing parameters: (a) Elevation, (b) Slope gradient, (c) Slope aspect.

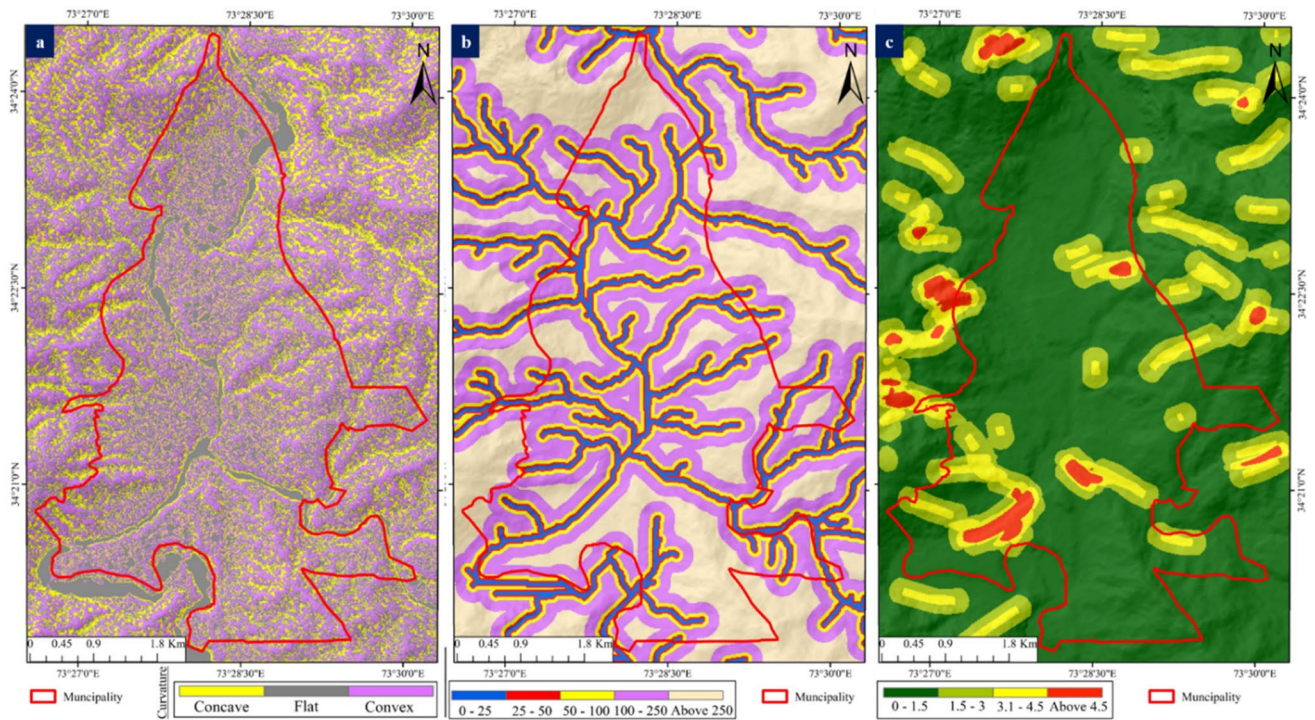


Figure 5. Groundwater potential influencing parameters: (a) Curvature, (b) Distance to streams, (c) Lineaments density.

as they illustrate how quickly rainwater percolates into the soil². The geological formations, soil absorption capacity, permeability, and slope of an area's geology all impact its drainage system⁴⁴. The drainage network of the study area was divided into five classes: 0 to 25 m, 25 to 50 m, 50 to 100 m, 100 to 250 m, and > 250 m, serving as evidence characteristics (Fig. 5b).

GW replenishment is notably influenced by surface and subsurface geological parameters, including lithology and geological structure⁴⁵. The area's lithology is a crucial factor in determining GW potential zones due to its influence on hydraulic conductivity⁴⁶. The lithological units were digitized by³⁷ using ArcGIS 10.8 (Fig. 2). Lineaments, recognized as curvilinear features on satellite images, play a significant role in hydrogeology by indicating permeable zones and GW transport pathways⁴⁷. The study extracted lineaments using Landsat-8 satellite imagery, which were then reclassified into four classes (Fig. 5c). The satellite image processing workflow involves both pre-processing and post-processing. In the preprocessing, the acquired satellite images undergo pre-processing, which includes atmospheric corrections. This correction is achieved using a radiometric calibration tool, ensuring that atmospheric interference is minimized or removed from the images. Post-processing involves applying a smoothness operation to the satellite images. The purpose of this operation is to reduce noise and enhance the visual quality of the images. One of the most important ecological parameters for determining GW occurrence is land use and land cover (LU/LC). The changes in LU/LC have a direct impact on GW flow and also directly influence hydraulic conductivity⁴⁸. The LU/LC of the area was prepared using Exliss ENVI and satellite imagery from Sentinel-2 with a cell size of 10 m x 10 m and less than 2% cloud cover, acquired from the official source of Copernicus Open Access Hub-EU. The LU/LC was classified into five classes: Water Bodies, Green Land, Forest Land, Barren Land, and Urban Land (Fig. 6). In this work, both input data and output results are on a regional scale. The input data, collected from various regional sources, includes a DEM at 12.5 m resolution, Landsat imagery at 30 m resolution, and a geological map at 1:50,000. The final GW potential map is presented at a resolution of 30 m, ensuring regional relevance and applicability. Overall, this study integrated diverse data layers from topography, hydrology, geology, and ecology to compute the GW potential model, providing valuable insights into GW availability and quality within the Muzaffarabad municipality.

Groundwater potential modeling methods

Weight of evidence (WOE)

The WOE is a bivariate statistical approach employed in many scientific investigations to assess environmental phenomena for almost six decades⁴⁹⁻⁵¹. Initially, this approach was created for the diagnostics and predictions of diseases⁴⁹. Then in the 1980s, this approach was employed for the probability assessment of minerals⁵². The WOE is based on the Bayesian statistical approach to evaluate the geospatial correlation of phenomena with their effective factors by assigning the weights to each part of the causative factor. This statistical modelling approach also assesses the positive (W^+) weight at the time of sample evidence happening to every specific variable that was predicted and the negative weight (W^-) that reflects the absence of the predicted variable at the time sample evidence not happening, and Contrast (C) which is a quantitative measurement of the relation among the class

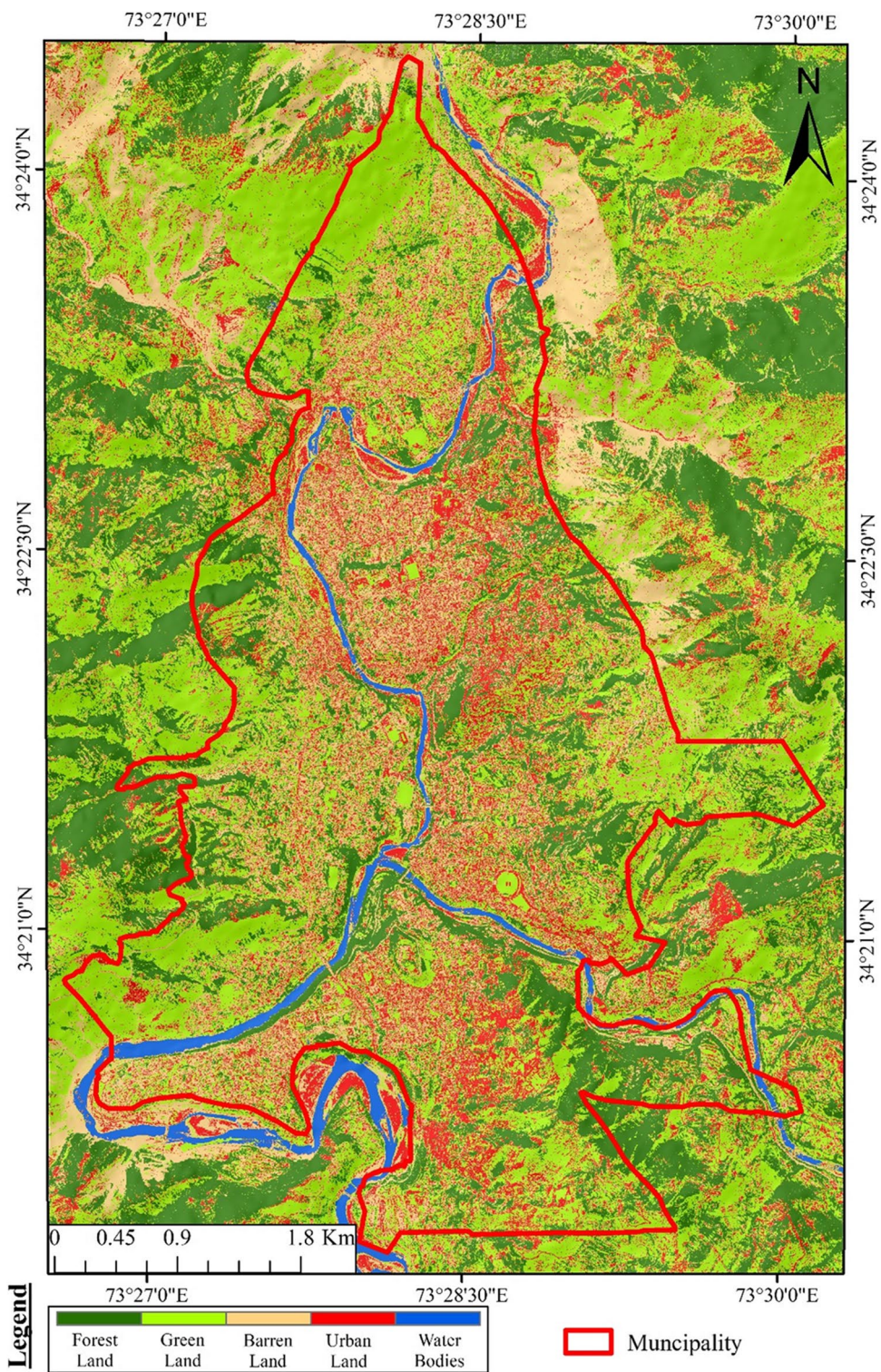


Figure 6. Landcover map of the Muzaffarabad municipality area.

of effective cause and evidence. The mathematical equation of the positive and negative weights is as follows in Eq. 1 and Eq. 2:

$$W_+ = \ln \left(\frac{\text{Happening in observed class}}{\frac{\text{Total happening area}}{\frac{\text{Non-happening area in observed class}}{\text{Total non-happening area}}}} \right) \quad (1)$$

$$W_- = \ln \left(\frac{\text{Happening area in other class}}{\frac{\text{Total happening area}}{\frac{\text{Total non-happening area in other class}}{\text{Total non-happening area}}}} \right) \quad (2)$$

and the contrast (C) value can be computed by the following formula (Eq. 3)

$$C = W_{\text{positive}} - W_{\text{negative}} \quad (3)$$

The WOE assessment evaluates the effectiveness of each influencing cause by correlating it with GW occurrence in different locations. Based on this correlation, specific weights are assigned to each influencing cause. These weights signify the relative importance of each cause in predicting GW potential. The WOE technique helps in determining the significance of different factors and their contributions to GW availability and quality in the study area. W^+ and W^- are the two measurements used to determine these weights. To receive a W^+ , the training points placed inside the class are evaluated. The W^+ score is larger than 0 if the class contains more training points than would be predicted by chance. The W^- is smaller than 0 if the number of training points outside the pattern is fewer than what would be predicted by chance. W^+ and W^- are combined to form contrast (C), a measure of the total relationship between a class and the training points. The weights of each influencing factor are determined by utilizing the indexes of W^+ and W^- . The probability model results from combining the weighted influencing parameter map into a probability map using the following mathematical summation (Eq. 4).

$$\text{WoE GW} = \sum W(1) + W(2) \dots + W(n) \quad (4)$$

In this study, WOE bivariate statistical analysis is performed using the ArcSDM (Spatial Data Modeller) which is an extension of the ArcGIS⁵². It is used to determine the weight and related statistics of influencing elements and to develop a posterior probability GW model.

Support vector machine (SVM)

Another popular classification approach based on statistical learning theory is the SVM, sometimes known as the maximum margin classifier. It was developed in the 1990s and is considered one of the most widely used approaches because of the performance efficiency in various algorithms. SVM is from the conceptual optimization premise and is fundamentally based on statistical learning theories. It is typically used to develop the highest possible generalization capabilities for machine learning's empirical relations and confidence intervals. The following Eq. 5 describes it.

$$K(x_a, x_a) = \exp \left(-\lambda \sum_{b=1}^n (x_{ab} - x_{atb})^2 \right) \quad (5)$$

Here in this eq. x_{ab} and x_{atb} are the i^{th} pair of observations of the predictor, n is the number of predictors, λ is a tuning parameter that accounts for the smoothness of the decision boundary, and K stands for the kernel function. Hyper-parameter tuning is a critical step in ML, involving systematic testing and optimization to improve model performance. We employed a Radial Basis Function (RBF) kernel, which is particularly effective for classification and regression tasks involving complex relationships. In this study, the best tuning parameters were determined using the grid search (GS) technique. The process began with training the SVM model with the RBF kernel by finding the optimal kernel parameters (cost (C) and gamma (γ)) using GS. After GS analysis, the optimum parameter values were identified: a gamma value of 0.3 and a cost value of 1 were used in our model. To save time and for quick result modelling an innovative tool pack named LSM (FE tool) developed by⁵³ is used in ArcGIS to generate the SVM-GW model.

Water quality analysis

Water quality analysis begins at the sampling site immediately after collecting the samples. The reliability of the analysis heavily relies on having the necessary facilities, tools, and equipment, as well as following recommended sampling procedures. During the sampling process, utmost care is taken to eliminate any external factors that could alter the composition of the samples and affect the results. To facilitate this, water samples for physical, biological, and chemical analysis were collected using 1-litre capacity polystyrene bottles. The location coordinates of each spring were recorded using the Global Positioning System (GPS).

At the sampling site, physical parameters such as Temperature, Electrical Conductivity (EC), Turbidity, Odour, Colour, and Taste were measured, and the GPS coordinates of each sampling site were recorded. All collected samples were placed in a chill box to maintain their integrity and then transferred to the laboratory for further

bacterial and chemical analysis. Advanced multi-parameter equipment was used to analyze parameters like pH, conductivity, temperature, and turbidity. For the chemical parameter analysis, a high-tech Spectrophotometer Lovibond XD-7500 was employed. The chemical parameters included total hardness, Ca, Mg, Chloride, Nitrite, Nitrate, and Sulphate. Bacterial test kits were also incubated for 48 h, after which they were carefully examined. Following the analysis, a table was created to record the results for each parameter based on the collected samples. The spatial distribution map of each parameter in the study area was developed using ArcMap 10.8. Each parameter was then reclassified into different classes, ranging from least to suitable, based on the standards set by the WHO for drinking water. In summary, the water quality analysis process is conducted meticulously, starting from the sampling site to the laboratory analysis and mapping, ensuring accurate and reliable results for each parameter in the study area.

Water quality index (WQI) modeling

The Water Quality Index (WQI) is an effective way to summarize all quality parameters into a single value, expressing the suitability of water resources for human consumption. As described by⁵⁴, the WQI is a comprehensive ranking that considers the combined impact of multiple factors affecting water quality. The WQI is determined by assessing the cumulative effect of both human-induced and naturally occurring activities, based on specific parameters in the hydro-geometry properties of the water samples.

To calculate the WQI values for each sample location, the average concentration of determinants (TDS, EC, Cl^- , HCO_3^- , PO_4^{3-} , Ca^{2+} , NO_3^- , Na^+ , K^+ , and Mg^{2+}) is used for both dry and wet periods. The weighted Overlay method, commonly used for multi-criteria analysis such as site selection and appropriateness models, is employed to develop the WQI model. Each thematic layer is assigned a weight based on its significance using the Analytic Hierarchy Process (AHP) method. The developed thematic layers in raster format are integrated into ArcMap 10.8, and the model is constructed using the weighted overlay tool. Each reclassified thematic layer is then weighted according to its importance based on the WHO drinking water standards.

Results and discussion

Factor analysis

The slope of an area has a significant impact on both GW potential as well as GW infiltration. Table 1 illustrates the significance of slope as a key topographic factor influencing flat-mild slope areas indicating their more suitability for GW presence in the study area. The GW potentiality is greater in flat and gently sloping terrains⁵⁵. In the flatter slope terrain area, there is a higher probability that GW can accumulate.

The topographic factors are also essential to determine the elevation of the GW⁵⁶. The spatial significance of the slope's aspect as an effective topographic element is shown in Table 1. The southwest direction of the slope in the area has the maximum potential of GW in the study area based on the maximum weight value of WOE analysis. This is because the slope in the southwest direction in the study area receives more rainfall and moisture than in other directions. Table 1 shows the WOE spatial analysis of elevation for the GW. Based on the analysis study reveals those areas with an elevation range of 611–687 m have more potential for GW in the study area. Low-elevation areas are more susceptible to runoff⁵⁷. The greater water absorption and recharging of the water are present in the low-elevation terrain plain area⁵⁸. The slope curvature is also analyzed as an effective topographic element to evaluate the potential of GW by using WOE, shown in Table 1. The analysis revealed that the slope in the study area with concave geometry has maximum GW potential. This category of curvature holds more water over a long period⁵⁷. Concave areas often act as recharge zones, collecting water runoff, while flat areas may indicate favourable conditions for aquifer development.

The proximity of drainage network analysis is important for delineating the GW in an area. The characteristics of surface and subsurface formation are reflected in the drainage pattern. The WOE assessment spatial relationship of drainage proximity with GW is shown in Table 1. The study concluded that an area of up to 50 m of drainage has the maximum potential for groundwater in the study area. Lithology as an effective geological element and its spatial significance are shown in Table 1. The WOE analysis concluded that the surficial deposit of the study area has the potential of GW as it has the maximum weight value among other lithological units in the study area. This is because most of Muzaffarabad's municipality area comprises surficial deposits. The predominance of surficial deposits in the study area, coupled with its low elevation, suggests a high potential for GW due to the characteristics of these lithological units. This aligns with findings from⁵⁹, highlighting how such geological conditions can significantly influence GW availability and potential in Muzaffarabad's municipality area. In hard rock terrains, lineaments play a significant role in GW recharge; adjacent to the lineaments zone, the GW occurrence potential is high⁶⁰. The lineament intersections are considered good GW potential zones. The spatial relation of GW and the density of lineaments in the study area are analyzed (Table 1). The analysis concluded that the class range of above 4.5 of evidential parameter lineament density has more potential for GW in the study area as lineaments provide pathways for groundwater movements. The GW potential is high near high-density lineament zones and vice versa⁶¹. These results are in line with a recent study conducted by⁶² in Lahore, Pakistan. Land use/land cover is also an important effective factor for GW⁶³. Doke, et al.⁵⁹ observed that high-density lineament zones may not universally indicate very high GW potentials. This discrepancy could be attributed to additional factors such as lithology, soil cover, and gradient, which play crucial roles in influencing subsurface water recharge dynamics. Table 1 indicates the spatial significance of the land use/land cover as an ecologically effective element. The water bodies followed by barren areas have the maximum GW potential in the study area.

Ground water potential model

Planning and long-term development of a region depend heavily on having a better grasp of the possibilities of the GW. The management of GW sustainably depends on this kind of knowledge. A detailed GW resource

Causative Factor	Classes	W+	W-	Contrast (C)	Weight
Slope Gradient	0–10	2.42	-0.45	2.88	2.42
	11–19	0.39	-0.09	0.48	0.39
	20–26	0.06	-0.01	0.08	0.05
	27–33	-0.37	0.08	-0.45	-0.37
	34–46	-2.21	0.39	-2.61	-2.21
	Above 46	0	0	0	0.05
Slope Aspect	North	1.36	-0.24	1.61	1.36
	Northeast	0	0	0	-3.22
	East	-1.55	0.16	-1.72	-1.55
	Southeast	0.51	-0.07	0.58	0.51
	South	-1.92	0.26	-2.19	-1.92
	Southwest	2.08	-0.32	2.41	2.08
	West	0.97	-0.10	1.08	0.97
	Northwest	-1.31	0.11	-1.43	-1.31
Elevation	611–687	2.58	-0.65	3.23	2.58
	688–780	1.17	-0.26	1.44	1.17
	781–940	-0.72	0.21	-0.93	-0.72
	941–1,287	-2.63	0.72	-3.35	-2.63
	Above 1,287	0	0	0	-3.22
Curvature	Concave	1.58	-0.30	1.88	1.58
	Flat	1.57	-0.58	2.15	1.57
	Convex	-1.46	1.57	-3.04	-1.46
Distance to drainage	0 – 25	1.62	-0.17	1.80	1.62
	25–50	1.62	-0.17	1.80	1.62
	50–100	0.49	-0.08	0.57	0.49
	100–250	-0.12	0.07	-0.20	0.07
	Above 250	-1.93	0.68	-2.62	-1.93
Lithological Units	Muree Formation	-0.59	0.21	-0.81	-0.59
	Paleocene-Eocene Sequence	0.01	0	0.01	0.09
	Muzaffarabad Formation	-2.11	0.34	-2.45	-2.11
	Stream Channel Deposits	0.01	0	0.01	0.09
	Surficial Deposits	2.65	-0.73	3.39	2.65
	Hazara Formation	-0.75	0.22	-0.98	-0.75
	Terrace Deposit	0	0	0	0.09
Lineament Density	0–1.5	-0.19	0.50	-0.69	-0.19
	1.5–3	-0.29	0.06	-0.35	0.06
	3.1–4.5	-1.05	0.07	-1.13	-1.05
	Above 4.5	1.40	-0.12	1.53	1.40
LU/LC	Forest	0.10	-0.04	0.14	-0.04
	Green Land	-0.76	0.41	-1.18	-0.76
	Barren	0.86	-0.16	1.02	0.86
	Urban Land	0.41	-0.07	0.49	0.41
	Water Bodies	2.14	-0.03	2.17	2.14

Table 1. Cumulative weights analysis for the causative parameters.

evaluation is necessary since its availability fluctuates throughout time and place. Eight effective thematic maps were developed and analyzed to develop a WOE and SVM composite response for identifying GW potential zones in the municipality of the Muzaffarabad (Fig. 7a, b). Based on a categorization of natural breaks (known as “Natural breaks (Jenks)”), the study area has been divided into five categories: very poor, poor, moderate, good, and excellent potential zones of GW (Table 2). The results showed that 31.1% (16.21 km²) of the area had excellent GW potential based on the WOE model, whereas the SVM model showed that only 20.3% (10.59 km²) of the area fall in the excellent potential zone. According to the WoE model, the good GW potential zone is 6.9% (3.58 km²), while the SVM model is 12.9% (6.74 km²) in the good potential zone. In the WOE model, 6.1% (3.20 km²) had moderate GW potential, 8.9% (4.65 km²) had low GW potential, and 47% (24.47 km²) had very low GW potential, whereas in the SVM model, 25.7% (13.47 km²) moderate, 22.3% (11.61 km²) Low, and 18.7% (9.75 km²) very low GW potential. The differences observed between the SVM and WoE models occur mainly outside the Muzaffarabad municipality, particularly in regions with very low to moderate GW potential.

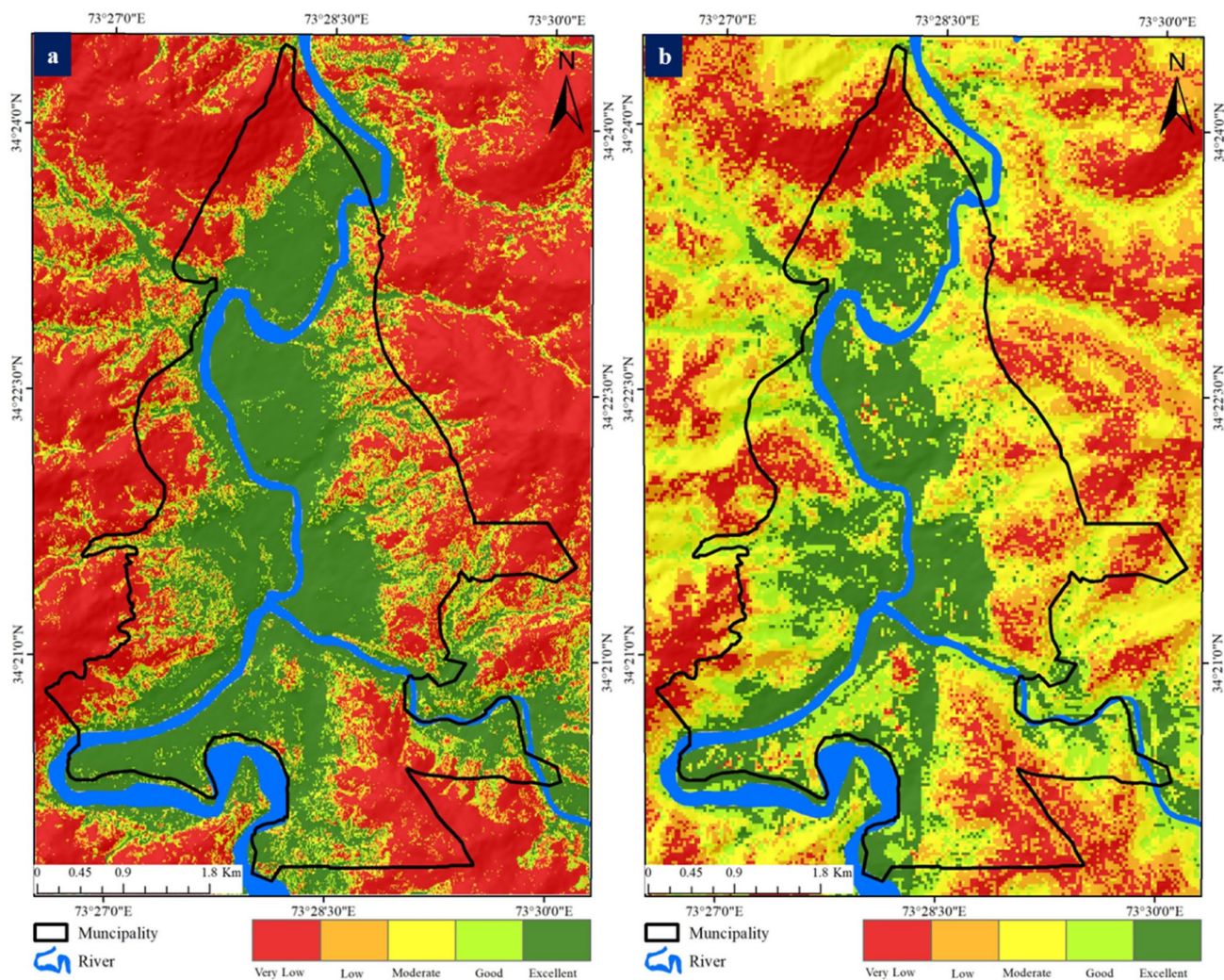


Figure 7. Groundwater potential potential map produced through a) WOE method, b) SVM model.

Sr. No	Class	Area (Km ²)		Area %age	
		SVM	WOE	SVM	WOE
1	Very Low	9.75	24.47	18.7	47.0
2	Low	11.61	4.65	22.3	8.9
3	Moderate	13.41	3.20	25.7	6.1
4	Good	6.74	3.58	12.9	6.9
5	Excellent	10.59	16.21	20.3	31.1

Table 2. The total area of GW potential zones using WOE and SVM model.

These differences are minimal in the core area of interest and do not significantly impact the study's findings. The variation can be attributed to the distinct methodologies: SVM is sensitive to kernel selection and parameter tuning, which influences its ability to model complex patterns, while WOE relies on the logistic transformation of input variables, providing a clear interpretation of the relationship between categorical and binary variables. Consequently, some degree of variation between the models is expected, particularly in areas outside the primary focus of our study. In the GW potential model, the influence of topographic conditions, particularly slope and elevation, is evident. Areas with high elevation and steep slopes, often composed of Pre-Cambrian and Cambrian formations, exhibit low GW potential due to the limited infiltration capacity. These formations are typically characterized by low permeability, which hinders the infiltration of rainwater and contributes to rapid surface runoff, reducing GW recharge. Conversely, areas with low elevation and gentle slopes have higher GW potential. These regions are often covered by quaternary deposits, which are more permeable and allow for greater infiltration. The nearly flat terrain in these areas slows down the runoff, providing ample time for

rainwater to percolate into the ground, thus enhancing groundwater recharge. Our results also indicate that areas with concave morphology exhibit higher GW potential, as this terrain naturally channels and retains surface water, enhancing infiltration and contributing to increased groundwater recharge. Our results demonstrate that GW potential is highest near drainage networks due to the increased availability of surface water, which promotes infiltration. The proximity to water bodies ensures a continuous supply of water, while the high lineament density enhances the permeability of the subsurface, allowing more water to percolate and recharge the aquifers. This combination of factors creates ideal conditions for sustaining higher groundwater levels in these areas.

In terms of area, Chella Bandi, Lower Plate, Domel, Lower Chatter, and Ambore exhibit higher GW potential compared to Tariqabad, Dhanni, and Mujajar Colony. The GW potential maps produced by both SVM and RF show that “high” and “very high” classes are concentrated in the watershed area, particularly around the river systems. This finding aligns with previous studies by³² and⁶⁴, which also observed higher GW potential near rivers. The maps indicate that GW potential decreases as distance from rivers and lineaments increases. Low-elevation areas depict a high potential for GW which is in line with the findings of²⁶. They found that elevation is the most influential factor for groundwater in the Markazi Province, Iran.

Validation of ground water potential model

To measure the accuracy of models, the “Area under the curve” (AUC) was calculated from SRC (training) and PRC (testing) datasets. The model performance is considered to be excellent if the AUC value is close to 1–0.9, < 0.9–0.8 good, < 0.8–0.7 medium, < 0.7–0.6 sufficient, < 0.6–0.5 bad and it would be taken as poor if its AUC value is < 0.5. In this study, 70% of GW inventory points from the study area are used to develop the model, while 30% remaining are used to validate the model. The model’s predictive performance is evaluated using the receiver operating characteristics (ROC) curves. Figure 8a, b shows the AUC of the model, which reveals that the model is in the range of the good class. In terms of AUC-ROC, SVM outperforms the WOE model having higher PRC and SRC values. To improve model performance and prevent overfitting, a feature selection process was adopted⁶⁵. Pearson’s Correlation Method was utilized to assess correlations among variables. This analysis was conducted using the Semi-Automatic Feature Selection module integrated within the ArcGIS Environment Toolbox. The correlation matrix plot facilitated the visual examination of correlation coefficients, allowing the manual removal of highly correlated features from the dataset. In addition, two different techniques, WOE and SVM, were employed to mitigate bias and overfitting.

Spatial distribution of water quality parameters

Bacterial contamination

A total of 39 water samples, including boreholes/wells and springs, were collected within the Municipality limits of Muzaffarabad City, PAK, for an in-depth analysis of GW quality. The primary focus of this study was to assess the water quality parameters and identify potential bacterial and chemical contamination. The water samples were subjected to thorough analysis to evaluate the presence of bacterial contamination. Surprisingly, 30 out of the 39 spring water samples (76%) were found to be contaminated with bacteria (Table 3). Specifically, the presence of E.coli was observed in 30 samples, indicating a significant risk to public health.

Alarming results were observed regarding the suitability of the water for drinking purposes. Only 9 out of the 39 samples were deemed fit for consumption without the need for any treatment. This implies that a vast majority of the collected water samples require some form of treatment before they can be considered safe for human consumption. The presence of E. coli in drinking water samples, as highlighted by⁶⁶ in Taunsa Sharif, is a critical indicator of recent fecal contamination and poses significant health risks due to potential pathogenicity. The detection of E. coli in all analyzed samples indicates widespread contamination, suggesting inadequate sanitation or treatment practices in the water sources. The findings from this GW quality analysis emphasize

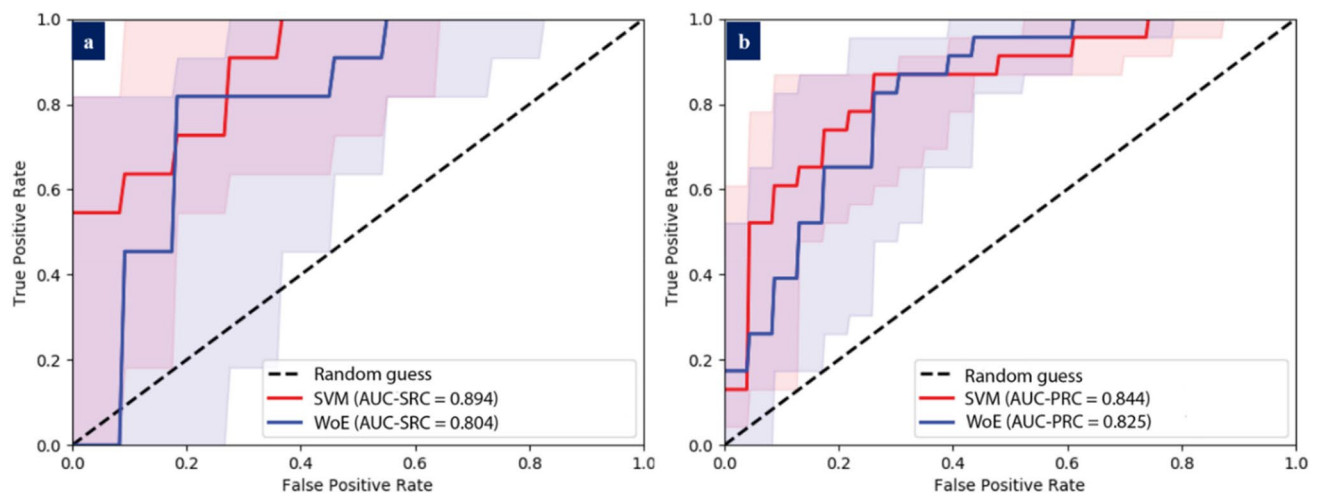


Figure 8. AUC-ROC curves: (a) SRC curve (b) PRC curve.

Variable	Units	Minimum	Maximum	Average	Standard Deviation	WHO standards (WHO 2022)
pH		6.2	8.4	7.03	0.6	6.5–8.5
EC	$\mu\text{S/cm}$	308	987	695.83	160.42	1500
Temp	C	10	27	18.75	5.68	Low temp is preferable
TDS	mg/L	80	582	326.64	149	1000
Turbidity	NTU	0	7.93	0.54	1.24	<5NTU
Ca	mg/L	17	207	99.61	43.7	200
Mg	mg/L	36	256	105.51	52.01	150
Hardness	mg/L	58	450	261.66	91.27	500
Chloride (Cl)	mg/L	9.5	221	35.77	35.51	250
Nitrite (NO_2^-)	mg/L	0.004	0.11	0.018	0.019	3
Nitrate (NO_3^-)	mg/L	0.04	40	15.065	10.89	50
Iron (Fe)	mg/L	0	0.1	0.009	0.02	0.3
Arsenic	mg/L	0	0	0	0	0.01
Sulphate (SO_4^{2-})	mg/L	16	403	142	90.49	250
Microbiological E-Coli	30 samples have +ive					-ive
Oduor	Un-objectionable					Un-objectionable
Colour	Un-objectionable					Un-objectionable
Taste	Un-objectionable					Un-objectionable

Table 3. Descriptive statistics of various GW parameters.

the importance of regularly monitoring water sources to ensure public health and safety. The high prevalence of bacterial contamination and chemical impurities underscores the need for proper water treatment and management strategies in the study area. Further investigation and implementation of appropriate remediation measures are essential to safeguard the well-being of the local population.

Physical parameters

Electrical conductivity (EC) The EC of water is influenced by a wide range of geological processes and human activities, such as ion exchange, reverse ion exchange, evaporation, silicate weathering, water-rock interactions, sulphate reduction, oxidation processes, and anthropogenic influences^{67,68}. The research findings, as shown in Table 3, indicate that the examined samples are in line with the standard set by the WHO (2022) for EC, which is 1,500 $\mu\text{S/cm}$. Our analysis found that all 39 collected samples exhibited an EC range below this limit, with the maximum value equal to or below 987, well within the acceptable range (Fig. 9a). Two samples lie in the range of 200–400; 8 samples lie in the range of 400–600; 18 samples lie in the range of 600–800 and 11 samples lie within the range of 800–1000 (Fig. 10a; Table 3).

Hydrogen ion concentration (pH) The pH measures the acidity or basicity of water, influenced by dissolved chemicals and the carbon dioxide-bicarbonate-carbonate equilibrium system. Water pH plays a crucial role in determining its usability for various purposes. The WHO's recommended pH range for most water sources is between 6.5 and 8.5, which is considered permissible. The presence of hydrogen ions in water is measured through the pH range, where a neutral pH indicates a balanced hydrogen ion concentration. In our current study, the pH range observed varied from 6.2 (minimum) to 8.4 (maximum), which falls well within the acceptable limit of 6.5 to 8.5, with an average pH of 7.03 (Fig. 9b; Table 3). Out of 39 samples, 20 samples lie within the pH range of 6–7; 12 lie within the range of 7–7.5; 4 lie within the range of 7.5–8 and 3 lie within the range of 8–8.5 (Fig. 10b).

Turbidity Turbidity in water is caused by suspended materials like clay, silt, organic particles, plankton, and microorganisms. It is a measure of light-scattering and light-absorbing characteristics. Out of 39 collected samples, only one bore well sample was found to be turbid and exceeded the permissible limits. The remaining samples from the springs and the bore well were within the WHO's guideline value for turbidity, which is 1 NTU (Fig. 9c; Table 3). Fahimah, et al.⁶⁹ analyzed the statistical relationship between topography and turbidity in Bandung regency, Indonesia and found that the concentration of turbidity was lower in the high topography area compared to the lower topography area.

Total dissolved solids (TDS) TDS refers to the combined presence of inorganic salts and low concentrations of organic matter in water, including major ions like carbonate, bicarbonate, chloride, sulphate, nitrate, sodium, potassium, calcium, and magnesium. TDS affects taste, hardness, corrosion properties, and encrustation tendencies. High TDS levels beyond the limits can lead to gastrointestinal irritation. The WHO recommended guideline value for TDS is 1000 mg/l. Our study found that most of the samples fall within this recommended range, with the minimum value being 85 while the maximum value is 582 mg/l (Fig. 9d; Table 3). Out of 39 samples, 9 samples lie within the range of 50–150 TDS value, 4 lies within the range of 150–300; 19 lies within the range of 300–450 and 7 samples lie within the TDS value of 450–600 (Fig. 10c). Rasheed, et al.⁷⁰ found that out of 82 GW samples, 74 samples of water are unsafe due to excessive TDS in District Jhelum, Upper Indus, Pakistan.

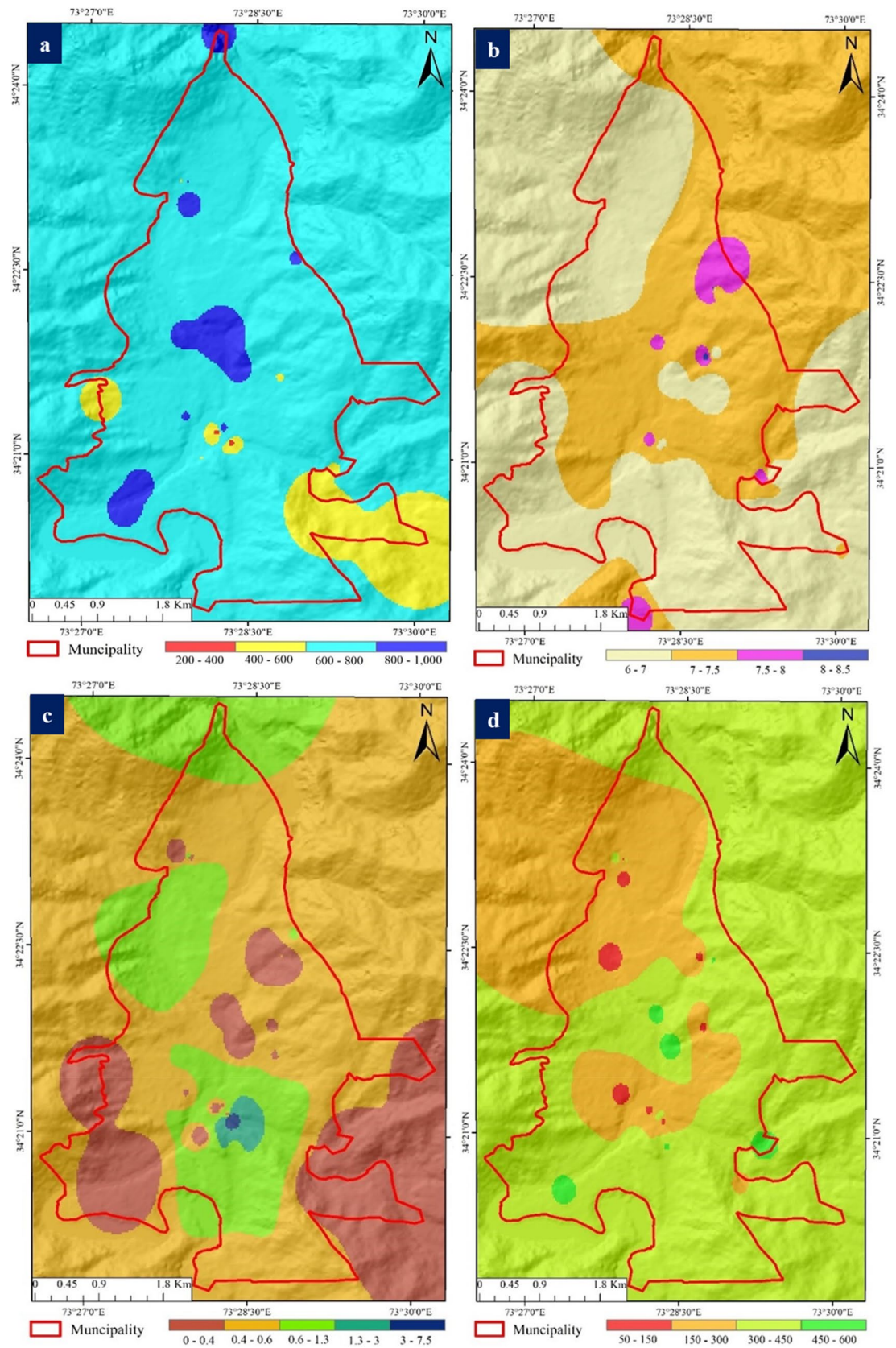


Figure 9. Spatial distribution maps: (a) Electric conductivity (b) pH (c) Turbidity (d) TDS.

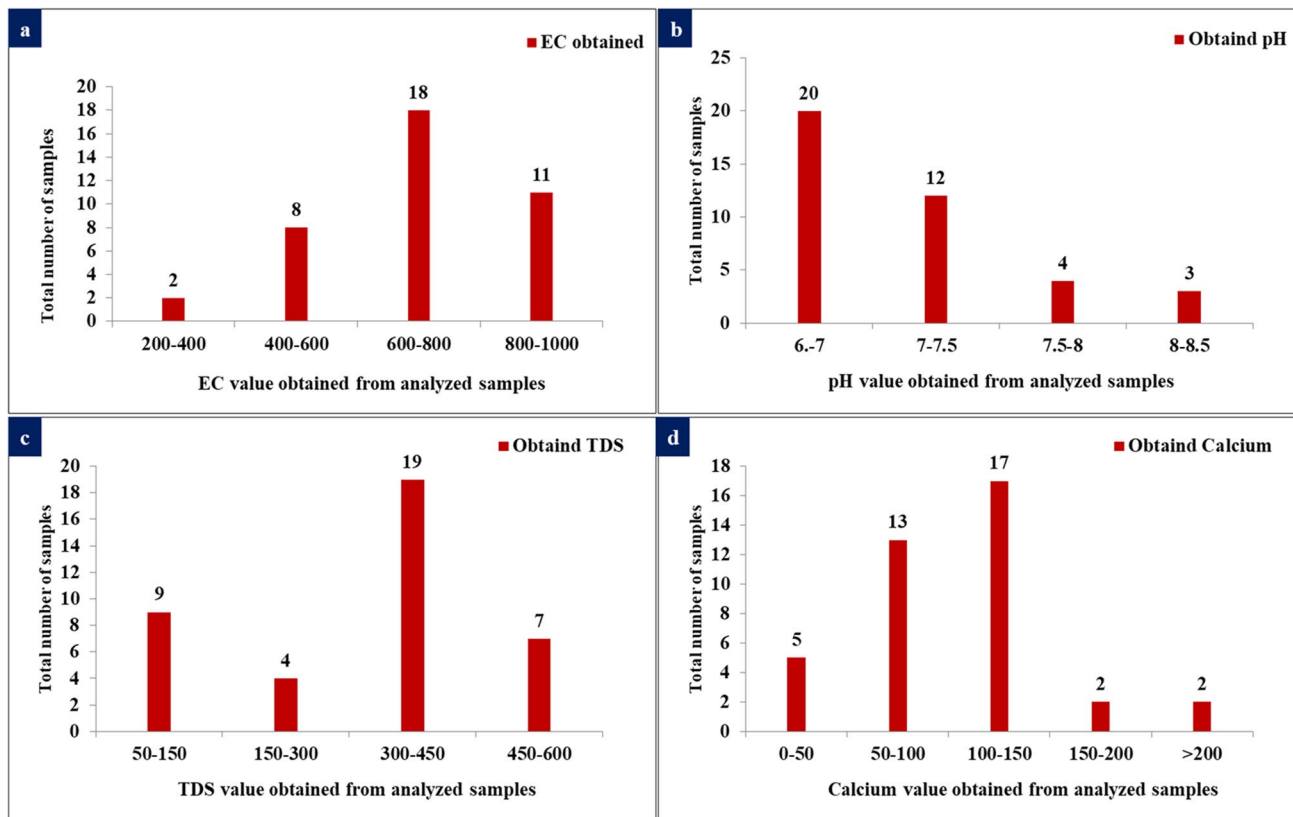


Figure 10. Spatial distribution graphs of various chemical tests: (a) EC (b) pH, (c) TDS (d) Calcium.

Chemical parameters

Among all 39 collected samples, 15 were chemically contaminated, while other water quality parameters were found within the permissible range. However, Turbidity in 1-Sample above the limit, Calcium in 2-Samples exceeds the permissible limits, magnesium in 7-samples exceeds the limits and sulphate in 10-samples exceeds the limits (Table 3).

Calcium (Ca²⁺) Calcium is the major abundant component in GW and contributes to water hardness when present along with magnesium. This mineral is crucial for various bodily processes, including blood clotting, nerve impulse transmission, and heart rhythm stabilization. The occurrence of calcium in GW is often attributed to deposits of limestone, dolomite, calcite, gypsum, and gypsiferous shale. The permissible range for calcium concentration is generally below 200 mg/l.

Among the 39 collected water samples, calcium contamination was observed in 2 samples, exceeding the WHO guideline value of 200 mg/l (Fig. 10d). Figure 11a indicates that the majority of samples fell within this range, except for two GW samples - one from a spring in Domail and another from a bore well in Sund Gali - where calcium levels were found to exceed the limit (Table 3).

Magnesium (Mg²⁺) Magnesium is another commonly found element in water and is also a significant cause of water hardness. The WHO guideline value for magnesium in water is 150 mg/l (Table 3). Among the 39 samples, only 6 exhibited elevated magnesium levels, while the majority of GW in the study area ranged between 100 and 150 mg/l, as shown in Figs. 11b and 12a.

Total hardness (TH) Water hardness is caused by the presence of elements such as calcium, magnesium, or ferrous (iron salts), such as chloride, sulphate, or bicarbonate ions. The terms “hard water” and “soft water” are commonly used to describe water with varying hardness levels. An acceptable compromise between corrosion and incrustation issues is typically achieved at a hardness level of approximately 100 mg of CaCO₃ per litre. However, WHO recommends a general guideline value of 500 mg/l for hardness due to its distinctive properties (Table 3). In the study area, the water hardness ranges observed during the survey fell within the permissible range, as illustrated in Figs. 11c and 12b (Table 3).

Chloride (cl) Chloride concentrations in GW tend to be higher than in surface water and can indicate the presence of organic pollutants. The high chloride levels in GW may result from factors such as lithological deposits, pollutant infiltration from sewerage systems, or seawater intrusion. Natural sources of chloride include NaCl,

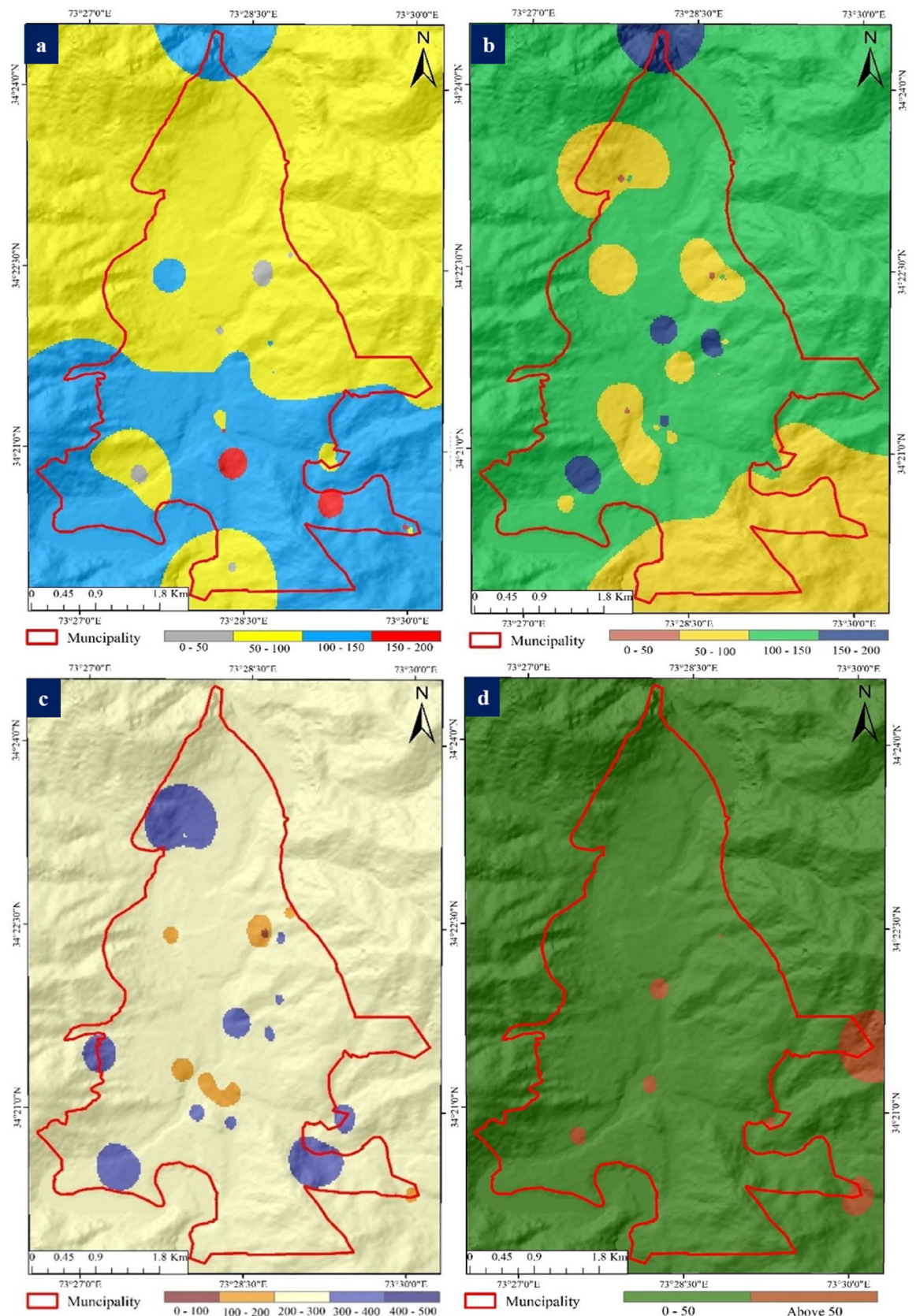


Figure 11. Spatial distribution maps: (a) Calcium (b) Magnesium (c) Total hardness (d) Chloride.

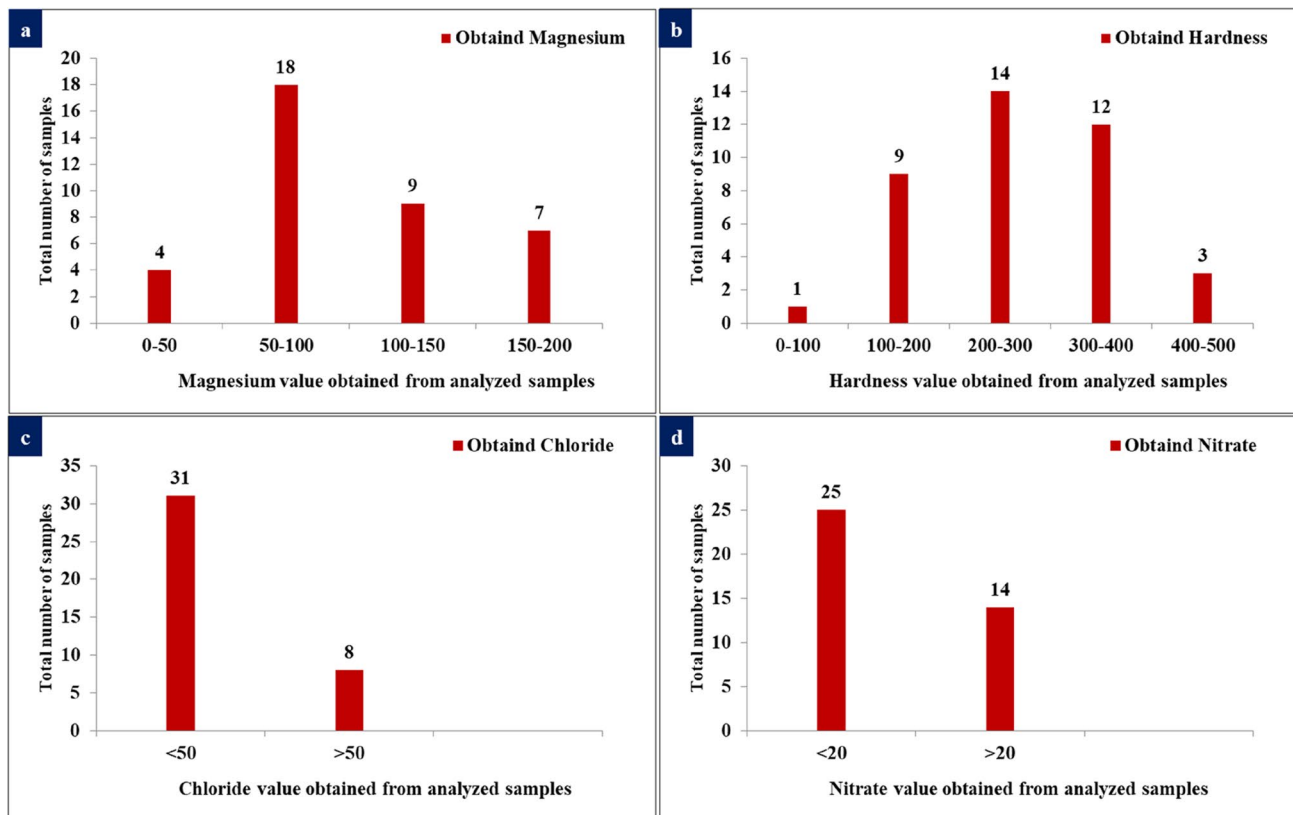


Figure 12. Spatial Distribution graphs of various chemical tests: (a) Magnesium (b) Hardness, (c) Chloride (d) Nitrate.

KCl, and Ca Cl₂ salts. The WHO guideline value for chloride in water is 250 mg/l (Table 3). In this study, all 39 samples analyzed showed chloride values below the WHO limit, as depicted in Figs. 11d and 12c.

Nitrate (NO₃⁻) Nitrate contamination in water is often attributed to various sources, such as fertilizer usage, decomposition of plant and animal waste, residential effluent, sewage sludge, industrial discharges, agricultural leachates, and climatic effects. Elevated nitrate levels in water can lead to health issues, including the “blue baby syndrome.” The WHO guideline value for nitrate in water is 50 mg/l (Table 3). The study area’s analyzed samples indicated nitrate levels below the WHO limit, as shown in Figs. 12d and 13a.

Sulphate (SO₄²⁻) Sulphate in groundwater can naturally originate from minerals like gypsum (CaSO₄ · 2H₂O), epsomite (MgSO₄ · 7H₂O), and barite (BaSO₄). Anthropogenic sources, such as waste discharge from mines, smelters, pulp and paper mills, textile mills, and tanneries, can also contribute to sulphate levels. The WHO guideline value for sulphate in water is 250 mg/l (Table 3). Out of 39 samples, 10 water samples exceed the WHO permissible limit (Figs. 13b and 14).

Groundwater quality index model

The groundwater index (GWI) model for the study area was developed by integrating the spatial distribution of physical and chemical properties of the groundwater samples. The GWI model, as depicted in Fig. 15, reveals that the majority of the locations exhibit moderate to good water quality, while some locations have excellent and some locations have poor water quality. The presence of bacterial contamination was detected in 30 samples, primarily attributed to poor solid waste disposal practices, open dumping, and inadequate sewage infrastructure leading to GW contamination. Alsalmeh, et al.⁷¹ investigated the GW quality for chemical and microbial contamination in Bhimber, (PAK) and found that almost all of the samples were grossly contaminated with *E. coli*. They also found that chlorite ion concentration is below the limits of WHO. Another study conducted by Khalid, et al.⁷² in Poonch, (PAK) underscores significant concerns regarding water quality. While chemical parameters generally met WHO standards, elevated lead levels pose a specific concern. However, the presence of biological contaminants, as indicated by positive results in coliform, total microbial load, and fungal tests, suggests widespread biological contamination. The improper disposal of liquid waste through open sewage and poorly designed septic tanks has contributed to the contamination issue⁷⁰. Additionally, the prevalence of open water sources, such as springs, increases the risk of contamination during the rainy season due to water runoff. To safeguard the GW quality and protect public health, addressing these contamination sources is imperative. Implementing proper solid waste management practices, improving sewage infrastructure, and promoting

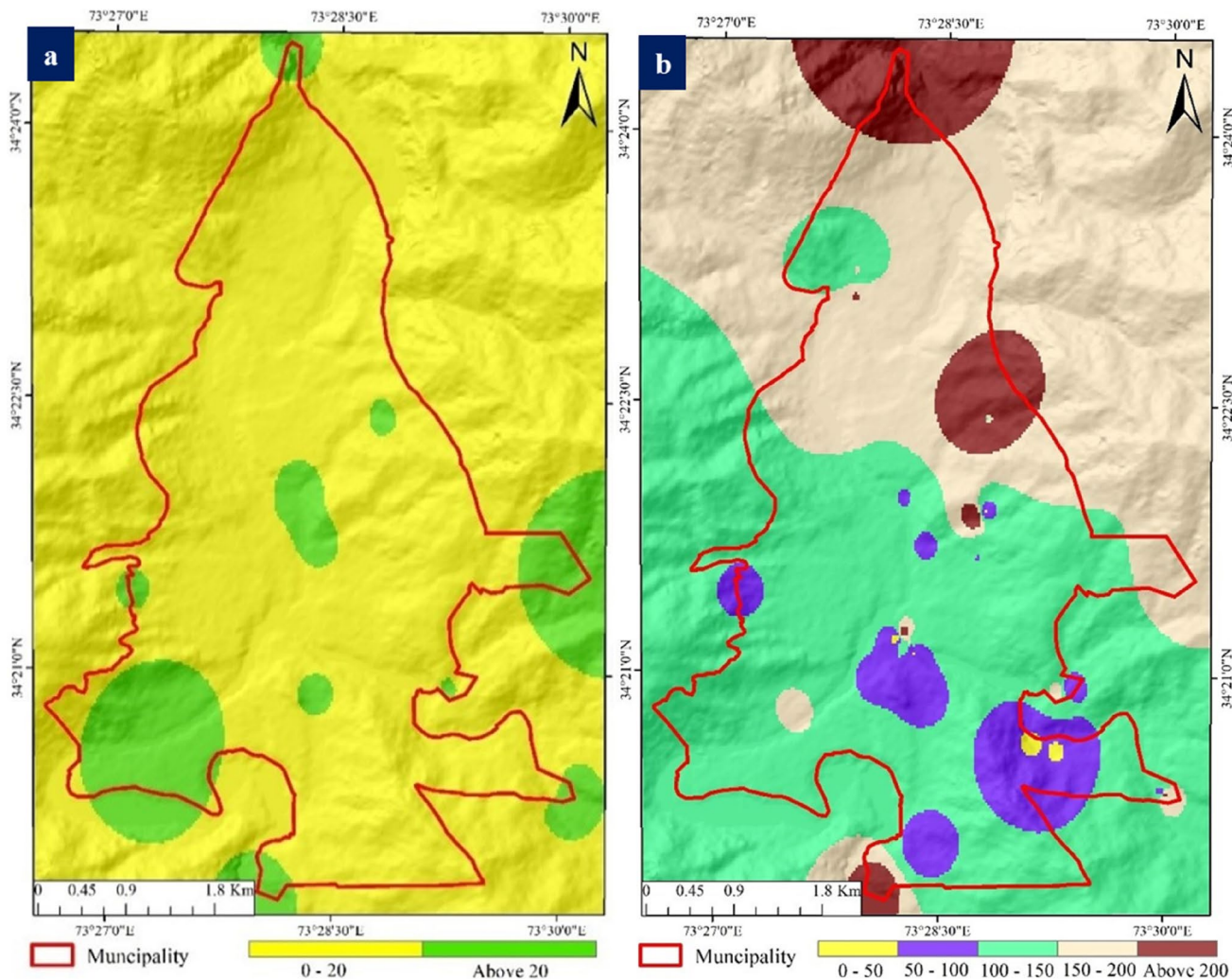


Figure 13. Spatial distribution maps: (a) Nitrate (b) Sulphate.

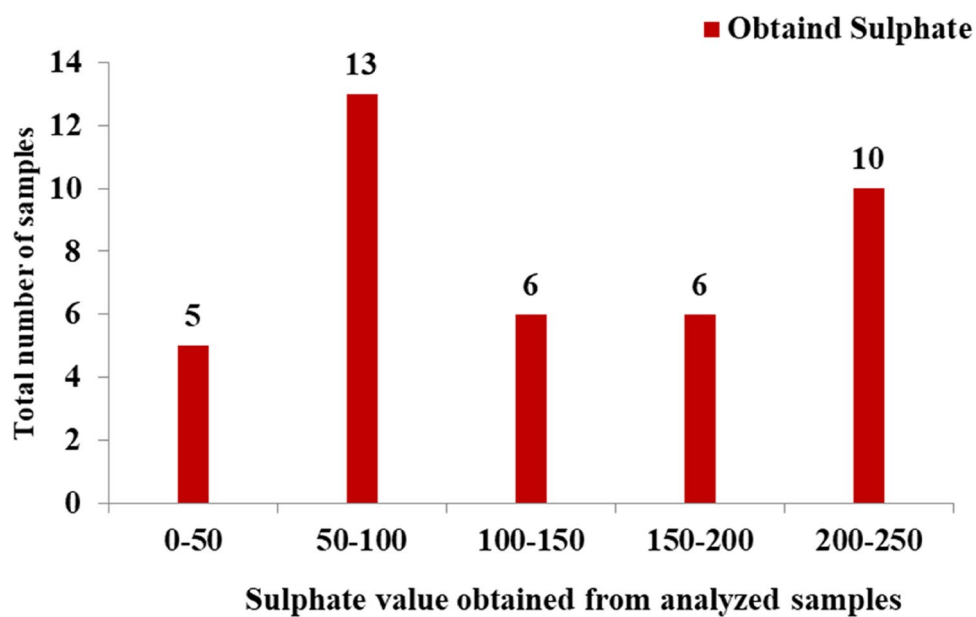


Figure 14. Spatial distribution graphs of sulphate.

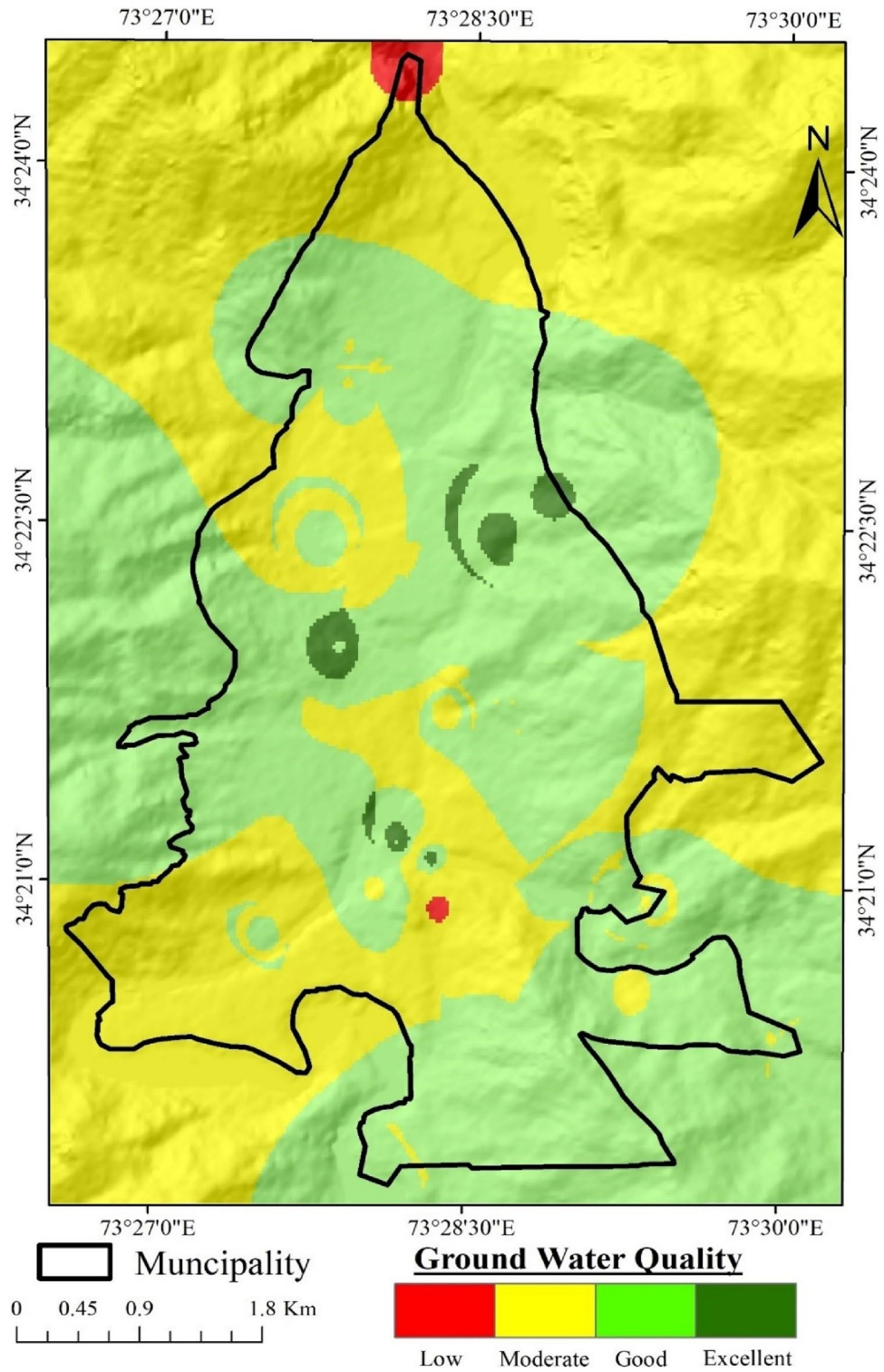


Figure 15. Groundwater quality index map of the Muzaffarabad municipality.

the use of closed water sources can significantly reduce the risk of bacterial contamination. Such measures will contribute to maintaining and improving the overall GW quality in the study area, ensuring a safe and sustainable water supply for the local population.

Conclusion

This study aimed to address the pressing water management issues faced by Muzaffarabad Municipality, Pakistan, through the integration of SVM and WOE techniques for GW potential mapping and water quality

assessment. The combination of these methods allowed for a comprehensive and accurate understanding of GW availability and quality in the region. GW potential mapping using SVM and WOE provided valuable predictive models, allowing the identification of potential GW zones. The weights assigned to each influencing factor in the WOE analysis helped in understanding the relative importance of different factors in predicting GW potential. Water quality analysis revealed the presence of various chemical parameters, some of which exceeded the WHO drinking water standards. The Water Quality Index was employed to summarize the water quality parameters, providing a comprehensive assessment of the suitability of water resources for human consumption. The study demonstrates that flat and gently sloping terrains, combined with low-elevation areas, exhibit higher GW potential due to enhanced water accumulation and absorption. The presence of concave slope geometries and proximity to drainage and high-density lineament zones further contribute to increased GW potential. However, some areas are classified as having very poor and poor GW potential, highlighting the need for targeted GW management strategies. Moreover, the assessment of GW quality indicates that bacterial contamination is a major concern, with substantial spring water and overall samples being contaminated with E.coli. This underscores the importance of implementing effective solid waste management practices and improving sewage infrastructure to prevent further contamination and safeguard public health. The study's findings can be used to guide sustainable GW management and conservation strategies in the Muzaffarabad municipality. Adequate measures to protect GW resources and improve water quality are vital to ensuring a safe and reliable water supply for the local population. Further research and continuous monitoring are recommended to track changes in GW potential and quality over time and assess the effectiveness of implemented management measures. By prioritizing efficient GW resource management, we can safeguard this precious resource for future generations and ensure the well-being and prosperity of communities worldwide.

Data availability

The data that support the findings of this study are available on request from the corresponding author.

Received: 13 November 2023; Accepted: 15 October 2024

Published online: 31 October 2024

References

- Duran-Encalada, J. A., Paucar-Caceres, A., Bandala, E. R. & Wright, G. H. The impact of global climate change on water quantity and quality: a system dynamics approach to the US–Mexican transborder region. *Eur. J. Oper. Res.* **256**, 567–581 (2017).
- Al-Ghrai, S. M., Razaq, I. B., Dwene, S. J., Ali, A. A. & Bajai, S. M. Evaluating phytomanagement as a biological reclamation method of salt-affected soils. *DYSONA-Appl. Sci.* **3** (1), 1–8 (2022).
- Wear, S. L., Acuña, V., McDonald, R. & Font, C. Sewage pollution, declining ecosystem health, and cross-sector collaboration. *Biol. Conserv.* **255**, 109010 (2021).
- La Vigna, F. Urban groundwater issues and resource management, and their roles in the resilience of cities. *Hydrogeol. J.* **30**, 1657–1683 (2022).
- Carrard, N., Foster, T. & Willetts, J. Groundwater as a source of drinking water in southeast Asia and the Pacific: a multi-country review of current reliance and resource concerns. *Water*. **11**, 1605 (2019).
- Chen, W. *et al.* Evaluating the usage of tree-based ensemble methods in groundwater spring potential mapping. *J. Hydrol.* **583**, 124602 (2020).
- Kalhor, K. & Emaminejad, N. Sustainable development in cities: studying the relationship between groundwater level and urbanization using remote sensing data. *Groundw. Sustainable Dev.* **9**, 100243 (2019).
- Pham, Q. B. *et al.* Groundwater level prediction using machine learning algorithms in a drought-prone area. *Neural Comput. Appl.* **34**, 10751–10773 (2022).
- Yadav, B., Gupta, P. K., Patidar, N. & Himanshu, S. K. Ensemble modelling framework for groundwater level prediction in urban areas of India. *Sci. Total Environ.* **712**, 135539 (2020).
- Swain, S., Taloor, A. K., Dhal, L., Sahoo, S. & Al-Ansari, N. Impact of climate change on groundwater hydrology: a comprehensive review and current status of the Indian hydrogeology. *Appl. Water Sci.* **12**, 120 (2022).
- Pande, C. B. *et al.* Delineation of groundwater potential zones for sustainable development and planning using analytical hierarchy process (AHP), and MIF techniques. *Appl. Water Sci.* **11**, 186 (2021).
- Ahmad, S. *et al.* Impact of water insecurity amidst endemic and pandemic in Pakistan: two tales unsolved. *Annals Med. Surg.* **81**, 104350 (2022).
- Ishaque, W., Mukhtar, M. & Tanvir, R. Pakistan's water resource management: ensuring water security for sustainable development. *Front. Environ. Sci.* **11**, 1096747 (2023).
- Daud, M. K. *et al.* Drinking water quality status and contamination in Pakistan. *Biomed. Res. Int.* **2017**, 7908183 (2017).
- Briscoe, J., Qamar, U., Contijoch, M., Amir, P. & Blackmore, D. Pakistan's water economy: Running dry. *World Bank, Washington, DC* 3540(1) 7908183(2005).
- Chowdhury, A., Jha, M. K., Chowdhury, V. M. & Mal, B. C. Integrated remote sensing and GIS-based approach for assessing groundwater potential in West Medinipur district, West Bengal, India. *Int. J. Remote Sens.* **30**, 231–250 (2009).
- Malik, M. I., Bhat, M. S. & Najjar, S. A. Remote sensing and GIS based groundwater potential mapping for sustainable water resource management of Lidder catchment in Kashmir Valley, India. *J. Geol. Soc. India.* **87**, 716–726 (2016).
- Jha, M. K., Chowdhury, A., Chowdhury, V. M. & Peiffer, S. Groundwater management and development by integrated remote sensing and geographic information systems: prospects and constraints. *Water Resour. Manage.* **21**, 427–467 (2007).
- Machiwal, D., Jha, M. K. & Mal, B. C. Assessment of groundwater potential in a semi-arid region of India using remote sensing, GIS and MCDM techniques. *Water Resour. Manage.* **25**, 1359–1386 (2011).
- Senanayake, I. P., Dissanayake, D., Mayadunna, B. B. & Weerasekera, W. L. An approach to delineate groundwater recharge potential sites in Ambalantota, Sri Lanka using GIS techniques. *Geosci. Front.* **7**, 115–124 (2016).
- Doke, A. B., Zolekar, R. B., Patel, H. & Das, S. Geospatial mapping of groundwater potential zones using multi-criteria decision-making AHP approach in a hardrock basaltic terrain in India. *Ecol. Ind.* **127**, 107685 (2021).
- Hasanuzzaman, M., Mandal, M. H., Hasnine, M. & Shit, P. K. Groundwater potential mapping using multi-criteria decision, bivariate statistic and machine learning algorithms: evidence from Chota Nagpur Plateau, India. *Appl. Water Sci.* **12**, 58 (2022).
- Maity, B., Mallick, S. K., Das, P. & Rudra, S. Comparative analysis of groundwater potentiality zone using fuzzy AHP, frequency ratio and bayesian weights of evidence methods. *Appl. Water Sci.* **12**, 63 (2022).
- Di Nunno, F., Granata, F., Gargano, R. & de Marinis, G. Forecasting of extreme storm tide events using NARX neural network-based models. *Atmosphere*. **12**, 512 (2021).

25. Husna, N., Bari, S. H., Hussain, M. M., Ur-rahman, M. T. & Rahman, M. Ground water level prediction using artificial neural network. *Int. J. Hydrology Sci. Technol.* **6**, 371–381 (2016).
26. Anh, D. T. *et al.* Assessment of groundwater potential modeling using support vector machine optimization based on bayesian multi-objective hyperparameter algorithm. *Appl. Soft Comput.* **132**, 109848 (2023).
27. Schratz, P., Muenchow, J., Iturrutxa, E., Richter, J. & Brenning, A. Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecol. Model.* **406**, 109–120 (2019).
28. Jalalkamali, A. & Jalalkamali, N. Adaptive network-based fuzzy inference system-genetic algorithm models for prediction groundwater quality indices: a GIS-based analysis. *J. AI Data Min.* **6**, 439–445 (2018).
29. Rasool, U. *et al.* Mapping of groundwater productivity potential with machine learning algorithms: a case study in the provincial capital of Baluchistan. *Pakistan Chemosphere.* **303**, 135265 (2022).
30. Echogdali, F. Z. *et al.* Application of fuzzy logic and fractal modeling approach for groundwater potential mapping in semi-arid Akka basin, Southeast Morocco. *Sustainability.* **14**, 10205 (2022).
31. Choubin, B. & Rahmati, O. In *Water Engineering Modeling and Mathematic Tools*. 391–403 (Elsevier, 2021).
32. Choubin, B. *et al.* In *Spatial Modeling in GIS and R for Earth and Environmental Sciences*. 485–498 (Elsevier, 2019).
33. Mosavi, A. *et al.* Ensemble boosting and bagging based machine learning models for groundwater potential prediction. *Water Resour. Manage.* **35**, 23–37 (2021).
34. Saha, R. *et al.* Urban aquifer health assessment and its management for sustainable water supply: an innovative approach using machine learning techniques. *Groundw. Sustainable Dev.* **25**, 101130 (2024).
35. Behzad, M., Asghari, K., Eazi, M. & Palhang, M. Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Syst. Appl.* **36**, 7624–7629 (2009).
36. Liu, D., Mishra, A. K., Yu, Z., Lü, H. & Li, Y. Support vector machine and data assimilation framework for Groundwater Level forecasting using GRACE satellite data. *J. Hydrol.* **603**, 126929 (2021).
37. Riaz, M. T., Basharat, M., Hameed, N., Shafique, M. & Luo, J. A data-driven approach to landslide-susceptibility mapping in mountainous terrain: case study from the Northwest Himalayas, Pakistan. *Nat. Hazards Rev.* **19**, 05018007 (2018).
38. Mahamat, A. D. O. & Bounab, A. The use of explanatory statistics for mapping groundwater potential zones in a semiarid area: case of the Waddai province, eastern Chad. *J. Afr. Earth Sc.* **205**, 105012 (2023).
39. Bhat, S. U., Dar, S. A. & Hamid, A. A critical appraisal of the status and hydrogeochemical characteristics of freshwater springs in Kashmir Valley. *Sci. Rep.* **12**, 5817 (2022).
40. Javaid, S., Shah, S. G. S., Chaudhary, A. J. & Khan, M. H. Assessment of trace metal contamination of drinking water in the Pearl Valley, Azad Jammu and Kashmir. *Clean-Soil Air Water.* **36**, 216–221 (2008).
41. Fida, M., Li, P., Wang, Y., Alam, S. M. K. & Nsabimana, A. Water contamination and human health risks in Pakistan: a review. *Exposure Health.* **15**, 619–639 (2023).
42. Niaz, A. *et al.* Flood modelling and its impacts on groundwater vulnerability in sub-himalayan region of Pakistan: integration between HEC-RAS and geophysical techniques. *Geomatics Nat. Hazards Risk.* **14**, 2257360 (2023).
43. Riaz, M. T., Basharat, M., Brunetti, M. T. & Riaz, M. T. Semi-quantitative landslide risk assessment of district Muzaffarabad, northwestern Himalayas, Pakistan. *Stoch. Env. Res. Risk Assess.* **37**(9), 3551–3570 (2023).
44. Dinesh Kumar, P. K., Gopinath, G. & Seralathan, P. Application of remote sensing and GIS for the demarcation of groundwater potential zones of a river basin in Kerala, southwest coast of India. *Int. J. Remote Sens.* **28**, 5583–5601 (2007).
45. Chenini, I., Mammou, A. B. & El May, M. Groundwater recharge zone mapping using GIS-based multi-criteria analysis: a case study in Central Tunisia (Maknassy Basin). *Water Resour. Manage.* **24**, 921–939 (2010).
46. Ayazi, M. H. *et al.* Disasters and risk reduction in groundwater: Zagros Mountain Southwest Iran using geoinformatics techniques. *Disaster Adv.* **3**, 51–57 (2010).
47. Sander, P. Lineaments in groundwater exploration: a review of applications and limitations. *Hydrogeol. J.* **15**, 71–74 (2007).
48. Quan, Q., Liang, W., Yan, D. & Lei, J. Influences of joint action of natural and social factors on atmospheric process of hydrological cycle in Inner Mongolia, China. *Urban Clim.* **41**, 101043 (2022).
49. Lusted, L. B. General problems in medical decision making with comments on ROC analysis. In *Seminars in nuclear medicine*. *WB Saunders.* **8**(4) 299–306 (1978).
50. Spiegelhalter, D. J. & Knill-Jones, R. P. Statistical and knowledge-based approaches to clinical decision-support systems, with an application in gastroenterology. *J. Royal Stat. Society: Ser. A.* **147**, 35–58 (1984).
51. Weed, D. L. Weight of evidence: a review of concept and methods. *Risk Analysis: Int. J.* **25**, 1545–1557 (2005).
52. Bonham-Carter, G. F. Weights of evidence modeling: a new approach to mapping mineral potential. *Stat. Appl. in the earth. Sci.* **171**–183 (1989).
53. Sahin, E. K. Implementation of free and open-source semi-automatic feature engineering tool in landslide susceptibility mapping using the machine-learning algorithms RF, SVM, and XGBoost. *Stoch. Env. Res. Risk Assess.* **37**, 1067–1092 (2023).
54. Sener, E. & Davraz, A. Assessment of groundwater vulnerability based on a modified DRASTIC model, GIS and an analytic hierarchy process(AHP) method: the case of Egirdir Lake basin(Isparta, Turkey). *Hydrogeol. J.* **21**, 701–714 (2013).
55. Subba Rao, N. Seasonal variation of groundwater quality in a part of Guntur District, Andhra Pradesh, India. *Environ. Geol.* **49**, 413–429 (2006).
56. Sener, E., Davraz, A. & Ozcelik, M. An integration of GIS and remote sensing in groundwater investigations: a case study in Burdur, Turkey. *Hydrogeol. J.* **13**, 826–834 (2005).
57. Moghaddam, D. D., Rezaei, M., Pourghasemi, H. R., Pourtaghie, Z. S. & Pradhan, B. Groundwater spring potential mapping using bivariate statistical model and GIS in the Taleghan watershed, Iran. *Arab. J. Geosci.* **8**, 913–929 (2015).
58. Thapa, R., Gupta, S., Guin, S. & Kaur, H. Assessment of groundwater potential zones using multi-influencing factor (MIF) and GIS: a case study from Birbhum district, West Bengal. *Appl. Water Sci.* **7**, 4117–4131 (2017).
59. Doke, A., Pardeshi, S. D. & Das, S. Drainage morphometry and groundwater potential mapping: application of geoinformatics with frequency ratio and influencing factor approaches. *Environ. Earth Sci.* **79**, 393 (2020).
60. Srivastava, P. K. & Bhattacharya, A. K. Groundwater assessment through an integrated approach using remote sensing, GIS and resistivity techniques: a case study from a hard rock terrain. *Int. J. Remote Sens.* **27**, 4599–4620 (2006).
61. Rashid, M., Lone, M. A. & Ahmed, S. Integrating geospatial and ground geophysical information as guidelines for groundwater potential zones in hard rock terrains of south India. *Environ. Monit. Assess.* **184**, 4829–4839 (2012).
62. Khushi, M. *et al.* Delineation of groundwater potential zones with Analytic Hierarchy process based geospatial modelling approach in metropolitan expanse. *Desalination Water Treat.* **315**, 399–412 (2023).
63. Siddik, M. S. *et al.* The impact of land use and land cover change on groundwater recharge in northwestern Bangladesh. *J. Environ. Manage.* **315**, 115130 (2022).
64. Rahmati, O. & Melesse, A. M. Application of Dempster-Shafer theory, spatial analysis and remote sensing for groundwater potentiality and nitrate pollution analysis in the semi-arid region of Khuzestan, Iran. *Sci. Total Environ.* **568**, 1110–1123 (2016).
65. He, R., Liu, Y. & Zhang, H. Study on automatic classification of arrhythmias. *Feature Eng. Comput. Intell. ECG Monit.*, 113–141 (2020).
66. Javaid, M. *et al.* Bacteriological composition of groundwater and its role in human health. *J. King Saud University-Science.* **34**, 102128 (2022).
67. Yusuf, A. *et al.* Monitoring of emerging contaminants of concern in the aquatic environment: a review of studies showing the application of effect-based measures. *Anal. Methods.* **13**, 5120–5143 (2021).

68. Hao, C., Wang, Y., He, K. & Gui, H. Seasonal distribution of deep groundwater fluoride, geochemical factors and ecological risk for irrigation in the Shendong mining area, China. *Front. Environ. Sci.* **10**, 1024797 (2022).
69. Fahimah, N., Salami, I. R. S., Oginawati, K. & Thaher, Y. N. Variations of groundwater turbidity in the Bandung regency, Indonesia: from community-used water quality monitoring data. *HydroResearch*. **6**, 216–227 (2023).
70. Rasheed, H., Iqbal, N., Ashraf, M. & ul Hasan, F. Groundwater quality and availability assessment: a case study of District Jhelum in the Upper Indus, Pakistan. *Environ. Adv.* **7**, 100148 (2022).
71. Alsalmeh, A., Al-Zaqri, N., Ullah, R. & Yaqub, S. Approximation of ground water quality for microbial and chemical contamination. *Saudi J. Biol. Sci.* **28**, 1757–1762 (2021).
72. Khalid, S., Altaf, U., Altaf, U., Shah, R. & Parveen, G. Drinking water quality assessment of Union Council Dhamni, Poonch, Azad Jammu and Kashmir, Pakistan, using water quality index and multivariate statistical analysis. *Environ. Contaminants Reviews (ECR)*. **3**, 24–31 (2020).

Acknowledgements

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/332/45. The authors sincerely thank the Environmental Protection Agency (EPA) AJK for their vital support in providing chemical analysis of water samples. Special thanks to Arshad Iqbal for his crucial contribution to the data, which was key to this research.

Author contributions

M.T.R.I. and M.T.R. Conceptualization; M.T.R.I. M.T.R. A.R. methodology; M.T.R.I. M.T.R. A.R. writing—original draft preparation; M.T.R.I. M.T.R. A.R. A.A.B. J.M. H.G.A. writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through Large Research Project under grant number RGP2/332/45.

Declarations

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.G.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024